12-2011

MODELING AND SYNTHESIS OF FUNCTIONALIZED META-POLY(PHENYLENE ETHYNYLENE) HELICAL STRUCTURES

Ha Nguyen
Clemson University, hhnguye@g.clemson.edu

MODELING AND SYNTHESIS OF FUNCTIONALIZED META-POLY(PHENYLENE ETHYNYLENE) HELICAL STRUCTURES

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Chemical Engineering

---

by
Ha Hai Nguyen
December 2011

Accepted by:
David A. Bruce, Committee Chair
Mark C. Thies
Scott M. Husson
Karl R. Dieter

ABSTRACT

The ability of *meta*-poly(phenylene ethynylene) (mPPE) materials to undergo random coil to helix conformational changes under select conditions affords many opportunities for their use in sensor, separation, catalysis, and bio-related applications. Thus, to advance the development of these materials, a modeling procedure based on replica exchange molecular dynamics (REMD) simulation was developed to reliably assess factors affecting the folding behaviors of functionalized mPPE variants in solution. A combinational modeling study of 20 functionalized mPPEs in five solvent conditions provided insight into how mPPE secondary structure is impacted by the complex relationship between mPPE functional groups and solvent moieties. Further, these simulation results predicted mPPE structures that exactly matched those experimentally observed with eight previously studied mPPE systems. Using this REMD procedure, a series of new mPPE structures having an alternating arrangement of *exohelix* functional groups were developed, so as to optimize both the functionality and structure of the polymer. Seven new functionalized mPPEs were synthesized based on this concept. The folding behaviors of these new polymers were characterized using UV absorbance and fluorescence emission spectroscopy, and these results showed that the observed behaviors matched those predicted from simulation. Additionally, two mPPEs having both ester and amine functional groups were found to be soluble and form stable helical structures in water. The imine functional groups of these helical mPPEs were also used as metal ligands for the synthesis of a polymer supported manganese(salen) complex. Formation

of these catalytically active complexes demonstrated our ability to manipulate functional group interactions through the mPPE folding process. Additional modeling studies demonstrated that hydrogen bonds inside the helix cavity could be used as a highly effective stabilizer for mPPE helical structures. Finally, five mPPE helical structures having different *endohelix* functional groups were used as functionalized nanochannels in a modeling study that demonstrated how the through pore transmission of water can be controlled by the presence of specific *endohelix* functional groups. The ability to control this through channel flow is important for a variety of separation and bioactive species applications. The combined simulation and experimental efforts show that for the first time one can use computational methods to direct the synthesis of mPPE systems with specified folding behaviors.

DEDICATION

To my wife, Hai Le, and my family, for their love and support.

ACKNOWLEDGMENTS

To my advisor and mentor Dr. David A. Bruce, I would like to express my sincere gratitude for his dedicated guidance, support and encouragement during my graduate study, not only to my research but also to many aspects in the life of an international graduate student. My gratitude is also to Dr. James McAliley for his thoughtful discussions on theory and applications of molecular dynamics simulations, for the folding indicator scripts, and especially, for his invaluable time spent correcting and editing much of my writing.

I would like to express my appreciation to my dissertation committee, Dr. Karl Dieter, Dr. Mark Thies and Dr. Scott Husson for their time, advice and many valuable suggestions.

I also would like to thank Dr. Christopher Kitchens, Dr. Douglas Hirt, Dr. Amod Ogale, Dr. Dev Arya, Dr. Alex Kitaygorodskiy, Dr. Ken Marcus and Dr. Jason McNeil for their permission to use the analytical instruments in their laboratories; Kim Ivey for her help in collecting GPC and IR data; Jennifer Grimland, Nihar Ranjan, Sunil Kumar, Alfredo Picado and Carolyn Quarles in the Chemistry Department; Dr. Ward Burgess, Sourabh Kulkarni, Julian Velez, Mary Mitchell, Fiaz Mohammed and Sam Lukubira in the Chemical and Biomolecular Engineering Department for their time in helping me to use their equipment.

Many thanks are to go to all the former and current staff of the Chemical and Biomolecular Engineering Department: Diane Jackson, Terry McAllister, Joy Rodatz and Jay Lyn Martin for their constant support; and Mike Wilbanks and Bill Coburn for their help with countless things in the laboratory.

To the members of my research groups and to my friends in the department, Alma Iris Cordova, Bethany Carter, Ming He, Marlon Morales, Mary Mitchell, Fiaz Mohammed, Sam Lukubira, Daniel Wandera, Rebecca Alway-Cooper and Julian Velez, I would like to thank them all for their involvement and companionship. To the former and current undergraduate students, William (Trey) Batson III and Raymond (Ray) Smith, and international exchange student, Pan Papita, thanks for working with me. I also would like to thank all of the Vietnamese students in Clemson for their help and for the many things we have done together.

Financial support for this study is from National Science Foundation, award number CTS-0626143. I also would like to acknowledge the support of the staff from Clemson University's Cyber Infrastructure Technology Integration (CITI) group, for administration and maintenance of the computers used in this work.

TABLE OF CONTENTS

Page

Table of Contents (Continued)

Page

Table of Contents (Continued)

Page

Table of Contents (Continued)

Page

Table of Contents (Continued)

Page

Table of Contents (Continued)

Page

LIST OF TABLES

List of Tables (Continued)

Table Page

LIST OF FIGURES

List of Figures (Continued)

Table Page

List of Figures (Continued)

Table Page

List of Figures (Continued)

Table Page

List of Figures (Continued)

Table Page

List of Figures (Continued)

Table Page

List of Figures (Continued)

Table Page

List of Figures (Continued)

Table Page

INTRODUCTION

The ability to manipulate and control macromolecular structures at the molecular level often leads to interesting scientific findings and, additionally, the identification of novel applications for macromolecular chemistry. This vision becomes more practical thanks to the discovery of foldamers, a select class of artificial polymers that undergo a conformational transformation from random structures to highly ordered secondary structures. Originally designed as models to study isolated factors affecting the structures of biological macromolecules, i.e., proteins or DNA, these biomimetic foldamers create an opportunity to engineer and synthesize new macromolecules capable of performing specific functions, similar to their natural counterparts. The *meta*-poly(phenylene ethynylene) (mPPE) materials are a prime example of such foldamers, attracting attention because of their segmental structures and ease of synthesis.

Under suitable solvent conditions, certain types of functionalized mPPEs spontaneously coil into more compact, highly ordered helical conformations, making them ideal candidates in investigating the effect of π-stacking and solvophobic effects as drivers of these conformational changes. Their helical structures, similar to the alpha-helix motif in protein structures, are formed with six aromatic rings per turn. These structures are stabilized by π-stacking interactions between pairs of overlapping aromatic rings. Further, the solvophilic pendant groups, which ultimately reside in *exohelix* positions,

significantly limit unfavorable interactions between solvent molecules and the solvophobic backbone.

In general, functional groups can be attached at various positions on mPPEs, and are classified as *exohelix* or *endohelix* functional groups based on whether they are placed at exterior or interior positions on the polymer in its helical conformation. The *endohelix* and *exohelix* functional groups serve as the main parameters by which mPPEs can be highly customized using available synthesis routes, to form a variety of functionalized mPPE secondary structures. The ability to exploit this mPPE folding ability further allows accurate control over not only the functional group arrangements on the mPPE back bone but also their interactions with their chemical environments. Hence, this dissertation focuses on the study of mPPE materials and the use of functionalized mPPE helical structures in two new applications: firstly, as supports for metal containing catalysts and secondly, as nanochannels for selectively facilitating transport of water and other small molecules. On a broader scale, the results described herein clearly demonstrate the significance of using modeling results to guide the synthesis of novel self-assembled materials.

This dissertation contains six chapters divided into four main parts:

The first part of this dissertation (Chapter One) presents a brief literature review of mPPE materials, focusing on their synthesis, folding characteristics, modeling, and potential applications.

The second part of this dissertation (Chapter Two) describes modeling studies that were used to gain a better understanding of functionalized mPPE folding behaviors. These studies also provided a reliable method to assess the stable conformations of novel functionalized mPPEs. Given the limited availability of experimental data on mPPE materials, and the number of affecting factors, predicting the folding behaviors using only mPPE structural information and conventional chemical relationships, such as polarity, is not an easy task. Our successful modeling procedure, from which multiple situations involving folding behavior of functionalized mPPEs could be evaluated, offers invaluable information prior to experimental work with these polymers.

The third part of this dissertation (Chapters Three and Four) concentrates on the use of this modeling method, applying the aforementioned simulation protocol to the design of new helical mPPE structures that are functionalized with a variety of *exohelix* and *endohelix* functional groups. Subsequent experimental works present the synthesis and characterization of several new mPPEs that share a common structural motif. The model is further employed in an effort to produce a manganese(salen) complex using an imine functionalized mPPE helical structure as the complex ligand. This study illustrates the concept of using the mPPE folding phenomenon as an approach to precisely manipulate the functional group interactions so as to introduce chemically active sites into synthetic macromolecular structures.

The fourth part of this dissertation (Chapters Five and Six) explores the idea of using mPPEs to synthesize functionalized nanotubes. One modeling study is presented in which

we examine the possibility of using hydrogen bonds instead of π-stacking as the primary stabilizer of the mPPE helical structures, while another evaluates the extent to which *endohelix* functional groups ability can be used to control the helix cavity environment. These results laid the groundwork for a second modeling study, where we explored the concept of using mPPE helical structures as a platform for functionalized nanotubes that could be used as selective nanochannels for the transportation of water and other small molecules.

CHAPTER 1

LITERATURE REVIEW

The ester functionalized *meta*-poly(phenylene ethynylene) (mPPE) and its folding phenomena were first reported in 1997 by Moore and coworkers in an effort to develop a simple platform for studying the conformations of biomolecular macromolecules (Figure 1.1).[1-4] Extensive modeling and experimental work from their group[3, 5-22] and other researchers[23-40] have explored the synthesis and conformational characteristics of several functionalized mPPEs and brought insight to potential applications for these polymers. As each chapter in this dissertation has its own literature review, only a brief introduction to the various relevant aspects is presented here.

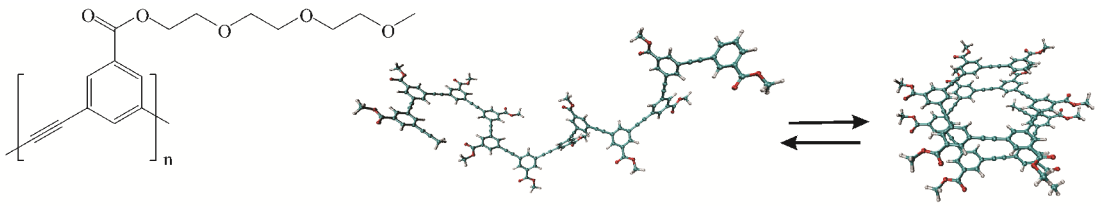


Figure 1.1 Structure of the ester functionalized mPPE reported by Moore and coworkers[3] and its folding reaction (side chains are removed for clarity). Atom colors for carbon, hydrogen and oxygen are teal, white and red, respectively.

1.1 mPPE folding phenomenon

The ester functionalized mPPE developed by Moore and coworkers[3] is a *meta* connected, conjugated aromatic backbone polymer with ether pendants attached to the ester functional groups. Under suitable solvent conditions, solvophobic driving forces cause mPPE structures of sufficient chain length to collapse into highly ordered, compact helical conformations that minimize the unfavorable interactions between their backbone and the solvent molecules. Figure 1.1 shows a representative folding reaction of the ester functionalized mPPE. The mPPE helical structures are similar to those of biological macromolecules, such as the double helix of DNA or the alpha helix of proteins, but in this case the helical structure is stabilized by π-stacking interactions between overlapping aromatic rings as opposed to the hydrogen bonding interactions that stabilize many biomolecule helices.[3-4, 41] Chain length dependence tests in acetonitrile revealed the minimum chain length, greater than eight repeat unit, required for an ester functionalized mPPE to exhibit a helical structure.[3, 13, 19] The helical conformation of the ester mPPEs are stable in a variety of solvents at moderate temperature, excluding chlorinated solvents and tetrahydrofuran.[3, 8] Experimental results for other functionalized mPPEs have revealed significant effects of the functional groups on the respective mPPE folding behaviors.[6, 10] Further, other interactions such as metal ion coordination,[14] the presence of chiral pendants[5, 7, 17, 40] or the binding between mPPE helical hosts and chiral guests,[16, 18] were also found to contribute to either the stability or the chiral bias of the mPPE helical structures.

The folding behavior and helical conformation of the ester functionalized mPPE as well as other mPPEs have been determined via a variety of analytical methods. The following section describes in greater detail those methods used for mPPE folding characterization.

1.2 mPPE folding characterization

Due to their aromatic conjugated backbone, the folding behaviors of functionalized mPPEs in solution can be directly characterized using a variety of structure sensitive analytical methods, namely UV absorbance, fluorescence emission, and, to a lesser extent, proton NMR, spin labeling NMR and X-ray scattering.[3, 9, 12, 15] These aforementioned spectroscopic techniques, particularly UV absorbance and fluorescence emission spectroscopy, have been used widely in the literature due to the ease of sample preparation and spectra interpretation. For example, the ratio between the intensity of UV absorbances at 303 nm and 289 nm for solutions containing the ester mPPE represents the ratio between *cisoid* and *transoid* conformations.[3] Because the helical conformation necessarily results in all ethynylene units arranged as *cisoid*, this ratio corresponds to the ratio between coiled and uncoiled mPPE structures in solution.[3] In chloroform or other solvents which do not promote organized folding, the ratio between those peaks is observed to be close to unity. On the other hand, in acetonitrile or other solvents that force mPPEs to adopt compact conformations (often helical), the ratio between *cisoid* and *transoid* conformations ranges from 0.5 to 0.8. Using fluorescence emission spectroscopy, one can also observe how changing the solvent from chloroform to

acetonitrile, for an adequate chain length mPPE, leads to a reduction in the intensity of the monomer emission peaks around 350 nm and a red shift in its emission wavelength. This quenching and shift are the result of an excimer effect of pairs of overlapping aromatic rings in close or proximity. UV absorbance and fluorescence emission spectra for a representative mPPE in acetonitrile and chloroform are shown in Figure 1.2 (top). These data show that changing the solvents leads to changes in the observed mPPE solution spectra, and these differences are the result of mPPE conformational transformation from *transoid* (in chloroform) to c*isoid* (in acetonitrile), as illustrated at the bottom of the figure. The presence of mPPE helical structures in suitable solvents can also be identified by the upfield shifting of the aromatic proton resonances in ^{1}H NMR spectra due to the π-stacking interactions that further shield these proton.[3] Additional evidence of the existence of highly ordered compact conformations was also reported from double spin labeling Electron Spin Resonance (ESR) NMR techniques.[11] The indicative line broadening effect caused by spin interactions was observed with systems where the solvent condition promotes the formation of mPPE helical conformations, as the spin labels are in proximity with each other, either directly above or in adjacent positions. From ESR-NMR results, the helix mPPE pitch was found to be consistent with a single helical turn containing six aromatic rings.[11] However, this technique is not widely used due to the requirement of preparing complex mPPE structures containing the active spin labeling groups.

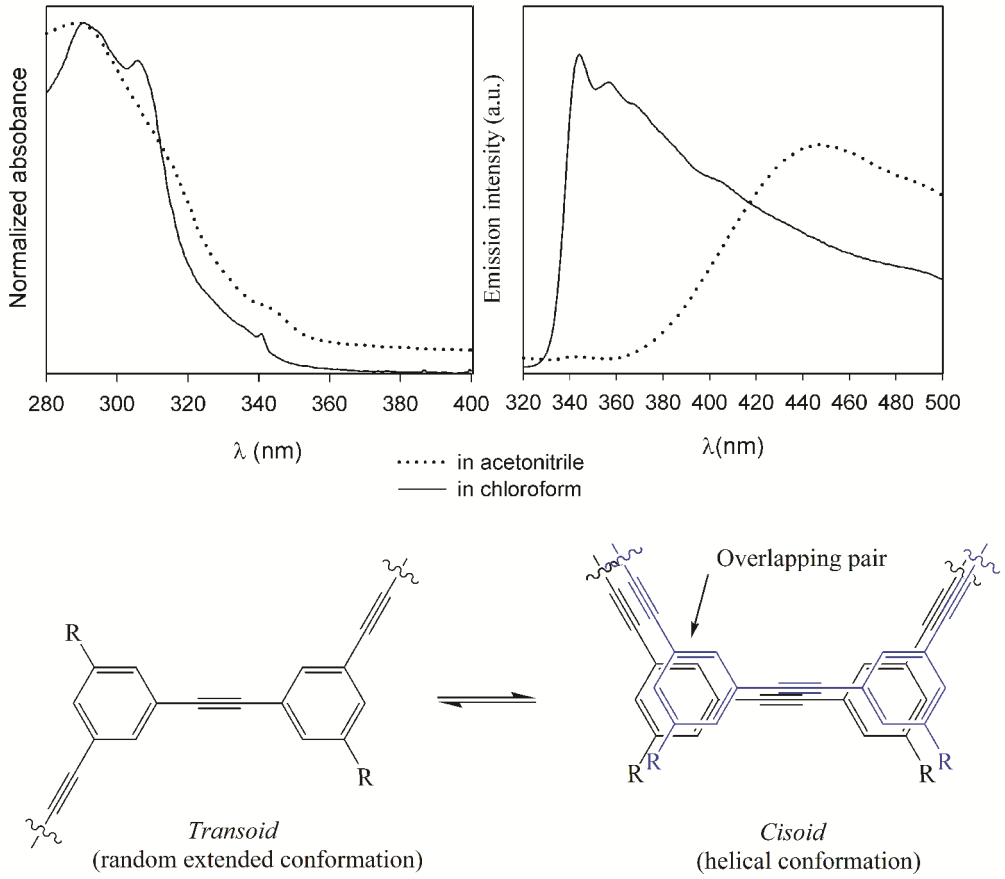


Figure 1.2 Representative UV absorbance (top left) and fluorescence emission spectra (top right) of an mPPE in acetonitrile and chloroform showing the variations in spectra that are resulted from solvent induced mPPE conformational transformation changes from a *transoid* (in chloroform) to c*isoid* (in acetonitrile) structure, as illustrated at the bottom.

To evaluate mPPE conformations in the solid state, small angle X-ray diffraction, wide angle X-ray diffraction and thin film UV absorbance have been used.[12] For X-ray diffraction, the mPPE was melted then annealed inside a capillary tube to increase the

resolution. Results indicate that the ester functionalized mPPE prefers lamellar packing or extended structures, however, having additional methyl functional groups at the *ortho* position to the ethynylene linkage forced it fold into helical structures that exhibited a columnar packing.

These previously published analytical methods, especially UV absorption and fluorescence emission spectroscopy, are important tools for studying mPPE secondary structure. In fact, these two spectroscopy techniques have been widely employed in many subsequent studies to assess the folding propensity of newly produced polymers, and are generally considered the standard methods for characterizing mPPE folding.

The following section provides a brief introduction to the experimental methods commonly used to synthesize functionalized mPPE materials.

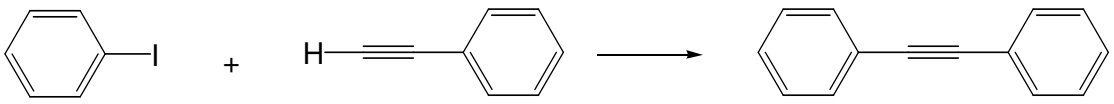
1.3 Synthesis of functionalized mPPEs

Functionalized mPPEs are routinely synthesized via the Sonogashira cross-coupling of halogenoarene and ethynylarene monomers using palladium catalysts.[3, 15, 42-44] The mechanism of this reaction includes four main steps: oxidative addition, *trans*-metalation, *trans/cis* isomerization and reductive elimination.[43-45] An example Sonogashira cross-coupling reaction between an iodoarene and an ethynylarene is shown in Figure 1.3.

Factors affecting this reaction include monomer composition, choice of base and co-solvents, the copper iodide concentration and the type of palladium catalysts.[15, 43-44]

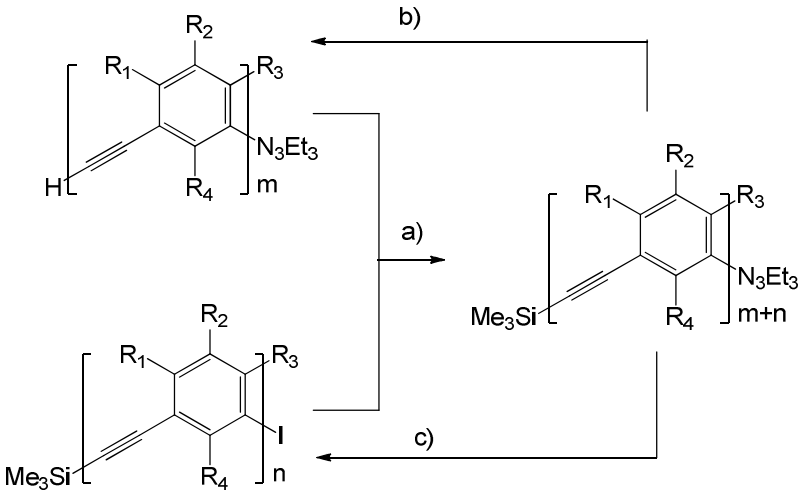
Iodoarenes are the most popular monomers for functionalized mPPE synthesis because of their greater reactivity and higher reaction yields.[15, 45] Normally, solvophilic functional groups are used as monomer pendants[3, 5, 46] to provide solubility to the polymer product, although there are reports that functionalized mPPEs have been synthesized without solvophilic pendants.[25-26] Several different approaches have been reported for the synthesis of mPPEs with a variety of functional groups (Figure 1.4 and Figure 1.5).

Moore and coworkers employed multiple step cross-coupling reactions followed by protection and deprotection of the halogeno and ethynyl functional groups to produce mPPEs having exact chain lengths from dimer up to 24-mer (Figure 1.4).[3, 7, 10, 15-17] This modulating method has an advantageously high degree of flexibility, allowing different types of functionalized monomers to be easily incorporated into the mPPE backbones.[6, 11, 47] However, it is not widely used due to the low overall yield of product that results from the high number of reaction steps required to synthesize a sufficiently long functionalized mPPE.



Reagents: $Pd(PPh_3)_4$, CuI, diisopropyl amine, toluene, 78 ^oC

Figure 1.3 Schematic Sonogashira cross-coupling reaction between an iodoarene and an ethynylarene (adapted from Ref. 33).

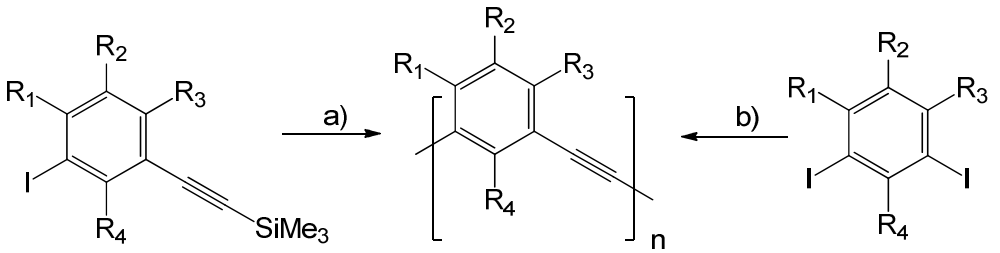


Reagents: a) $Pd_2(dba)_3$, CuI, PPh_3, Et_3N, CH_3CN, 70 °C
b) *tetra*-n-butylammonium fluoride, THF, r.t.
c) MeI, 110 °C

Figure 1.4 Synthesis routes employed by Moore and coworkers[3, 15] to synthesize functionalized mPPEs. The functional groups R_1, R_2, R_3 and R_4 could be different.

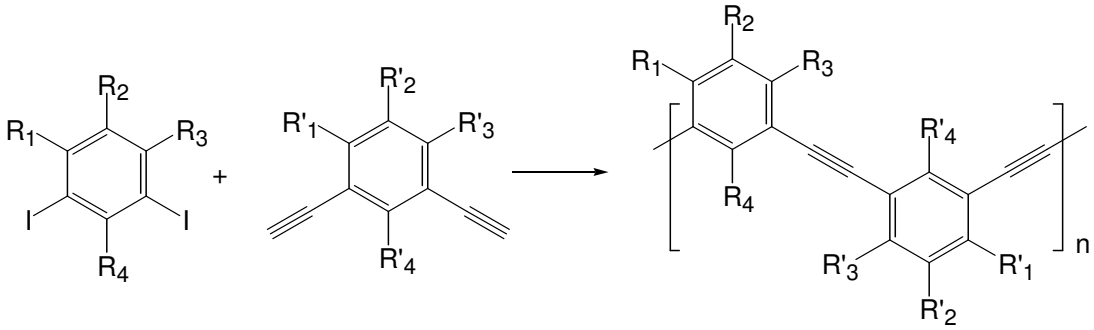
Several simpler one-step polymerization reactions have been introduced to obtain high molecular weight functionalized mPPEs (Figure 1.5). For example, a polymerization reaction between dihalogenoarene and diethynylarene monomers was reported by Arnt and Tew[25] and Huang et al.;[34] and between diiodoarene and acetylene by Li et al.[38] Additionally, Hecht and Khan[35, 45] proposed a synthesis route to reduce the diyne defect on the mPPE backbone. They used monomers containing both halogeno and ethynyl functional groups or using dihalogenoarene monomer and trimethylsilylacetylene with a small amount of water as the activator. These latter syntheses were conducted at room temperature and sometimes included the use of microwave heating/activation. Also, Zhao and Moore[47-48] developed a polymerization method that employed oligomers containing amide and aldehyde functional groups.

i)



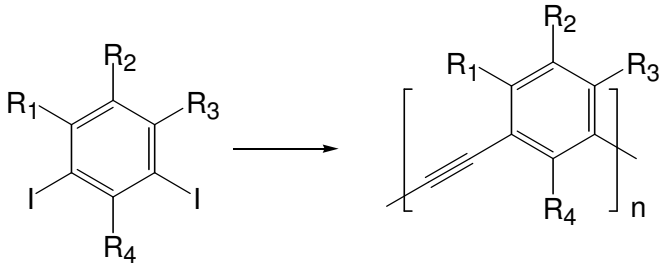
Reagents: a) [Pd(Ph$_3$)$_4$], CuI, H$_2$O, 1,8-Diazabicyclo[5.4.0]undec-7-ene (DBU), CH$_3$CN, r.t. or mw
b) Trimethylsilyl acetylene , [Pd(Ph$_3$)$_4$], CuI, H$_2$O, DBU, CH$_3$CN, r.t. or mw

ii)



Reagents: Pd(PPh$_3$)$_4$, CuI, diisopropyl amine, toluene, 78 ^{o}C

iii)



Reagents: C$_2$H$_2$, Pd(OAc)$_2$ + 4P(m-PhSO$_3$-Na+)$_3$, Et$_3$N, NaOH-H$_2$O, r.t.

Figure 1.5 Synthesis routes to functionalized mPPEs. From top, approaches employed by i) Hecht and Khan[32, 35], ii) Arnt and Tew[25-26] and iii) Li et al.[38] . The functional groups R_1, R_2, R_3 and R_4 could be different.

Those aforementioned synthesis routes for functionalized mPPEs, together with the previously described analytical methods for characterizing mPPE secondary structure

formation (helices formation), provide the necessary framework for our studies examining novel mPPEs and their applications. Several examples of those applications are presented in the following section.

1.4 Functionalized mPPE applications

Both functionalized mPPE helical and extended structures have been studied for potential applications. The first direct usage of mPPE materials was as simple biomolecule simulants, which could be used to quantify the effects of the π-stacking interaction between overlapping aromatic rings as well as the solvophobic effects on biomacromolecular structure.[3-4] Several functionalized mPPEs, including ester functionalized mPPEs, were synthesized for this purpose. Several other studies proposed applications that exploited the mPPE folding phenomenon and functionalized mPPE helical conformations. For example, mPPEs capable of hosting a comparable size guest molecule in the helix cavity[16] or able to form a silver complex with *endohelix* nitrile groups[14] indicated that mPPE helical structures could be used as molecular sensors. Hecht and Khan[32, 45] proposed an approach to synthesize organic nanotubes using an mPPE having ester functional groups and pendants containing C=C double bonds. UV radiation caused these double bonds to react and form covalent bonds between the pendants of two adjacent turns of the helical polymer, effectively locking the helical structures into nanotube forms. Tan et al.[39] presented an mPPE having electrolyte functional groups (SO_3-Na^+) capable of binding to a Ruthenium complex, changing the

complex from a non-luminescent state to a strongly luminescent state, which is similar to the binding effect found in DNA. A similar effect was also found in an mPPE having amine pendants reported by Huang et al.[34]

Besides these proposed applications, there were several studies exploiting only mPPE random polymer structures for their applications. An mPPE having both ammonium and ether functional groups alternating on two sides of its backbone was shown to have an interesting surface action and the ability to disrupt a phospholipid vesicle.[25-26] Further, a reversible hydrogel in acidic conditions using an acid functionalized mPPE was demonstrated by Li et al.[38] By changing the environment from low pH to high pH, the acid functional groups were converted into salts and the hydrogel was dissolved.

Our research group began using mPPE helical structures as a structural template for nanoscale mesoporous materials.[49] As a development from our early results, we have been conducting studies on using functionalized mPPE helical structures for other novel applications, in which the focus is on exploiting the folding reaction to control the functional groups interactions on the mPPE backbone. These materials are being actively explored for use as supports for catalysts, nano channels for biobased applications, chiral separating agents, and as molecular sensors.

1.5 mPPE modeling

Although extensive experimental work on mPPEs has provided considerable insight into the folding behavior of these materials, modeling studies offer the unique capability of examining the atomic level factors that contribute to folding bias of these polymers. Early efforts in mPPE modeling were conducted by Moore and coworkers, and they used Monte Carlo (MC) simulations, to search for local energy minima in the mPPE conformational space.[4] Their results suggest multiple possibilities for compact mPPE conformations, including the helical conformations. Subsequent modeling studies[23-24, 27-31, 37] were conducted by other research groups using more sophisticated molecular dynamics (MD) simulations for investigating multiple aspects of mPPE folding phenomenon, including the folding mechanism, the affecting factors and the structural variations. This section provides a brief review of the principles behind molecular dynamics simulations. More detailed information about these methods can be found in modeling texts authored by Allen and Tildesley,[50] Frenkel and Smith,[51] and van de Spoel et al.[52] as well as in mPPE modeling studies published by researchers at Clemson and elsewhere.[23-24, 27-31, 37, 49]

1.5.1 Molecular dynamics simulations[50-52]

The subject of all other published modeling studies on mPPE materials is the folding behaviors of functionalized mPPEs in solution, where the number of polymer atoms modeled are up to several thousand atoms. Those numbers of atoms are practically

outside the applicable range of most high accuracy quantum simulation techniques, given a reasonable time investment and currently available computing resources. Thus, all mPPE modeling studies to date have employed molecular simulations, mostly molecular dynamics.

The central idea behind MD simulations is to model the real time behavior of a system of *N* atoms by accurately representing how these atoms interact over time using simple algebraic equations. Knowing the trajectories of atoms over time enables one to calculate a variety of system properties (e.g., minimum energy conformations, heat capacity, phase equilibria, or diffusivity) using statistical thermodynamics and related transport concepts. In these simulations, the movements of the atoms are described via classical mechanics or more precisely Newton's equations of motion:

$$m_i \frac{\partial^2}{\partial t^2} r_i(t) = F_i \qquad i=1...N \tag{1.1}$$

where $r_i(t)$ is the 3-dimensional location at time t of atom i having mass m_i, and F_i is the momentary force acting on atom i.

The force F_i is defined as the derivative of the system potential energy function, V, which depends on the positions of all atoms in the studied system:

$$F_i = -\frac{\partial V}{\partial r_i} \qquad i = 1 ....N \tag{1.2}$$

The repeated numerical solution of Newton's equations of motion, enables one to calculate all atomic coordinates as a functional of simulated time, which yields the trajectories of the atoms in the studied system. The use of Newton's equations of motion (Eqn. 1.1) to describe atomic motion allows MD simulations to simulate biomolecules, polymers, and other large systems which could contain millions of atoms.

Generally, an MD simulation is begun by assigning atomic initial positions, initial velocities for each atom, and selecting an appropriate potential energy function. A simulation engine (or MD computer program) will use the initial specified conditions as a starting point to calculate the atomic forces acting upon each atom and integrate Newton's equations of motion over time. This integration process is repeated millions of times using small time increments that help ensure the fluid movement of atoms and avoid close atomic contacts or unrealistic bonding arrangements. There are other important parameters that must also be specified for the MD simulations to proceed; for example, the integration time step, the integration method, the development of neighbor lists, the temperature coupling and pressure coupling schemes. Figure 1.6 illustrates the four basic steps that are required for an MD simulation.

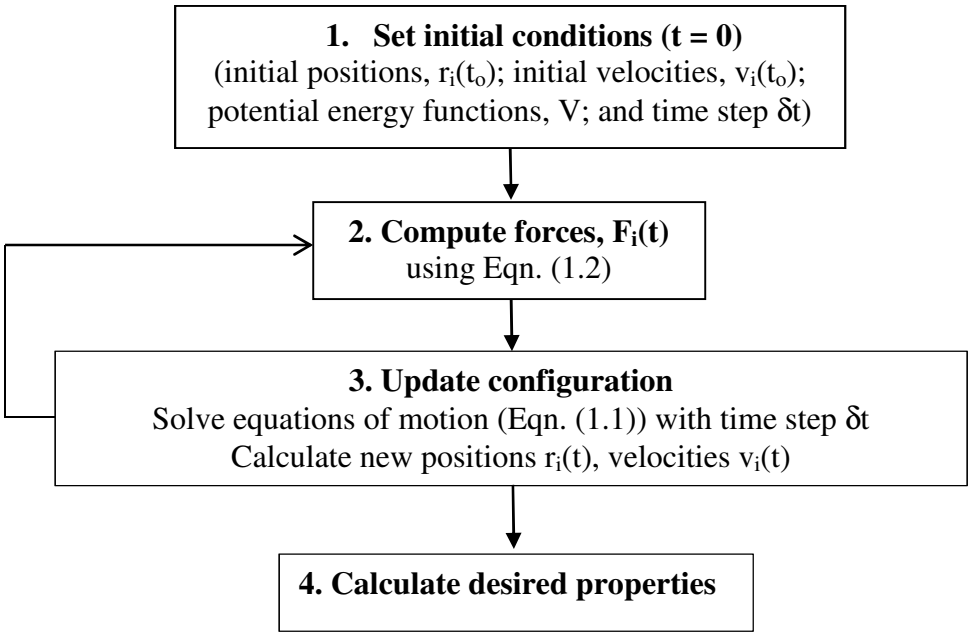


Figure 1.6. General computational steps involved with an MD simulation (adapted from Ref. 44).

The integration time step of an MD simulation, selected to ensure the stability of the simulated system, is often one femtosecond (10^{-15} s), which is a time less than the fastest hydrogen vibration frequency in the system (10^{-14} s). Given current computing resources, this small time step forces MD simulations to only be used for the processes having time scales from nanosecond (10^{-12} s) to microsecond (10^{-9} s), and depending on the number of atoms in the system, these simulations may take CPU-days or CPU-years to complete.

The potential energy function (V), which is more commonly called a force field, consists of an algebraic equation and set of atomic parameters that together determine the system

energy as a function of the atoms coordinates. This equation describes two types of atomic interactions: intramolecular (bonded) (e.g., bond-stretching, angle-bending, dihedral and cross-bonded terms) and intermolecular (non-bonded) interactions (e.g., electrostatic and van der Waals terms):

$$V = V_{bonded} + V_{non\text{-}bonded} \tag{1.3}$$

For example, several popular force fields in literature are OPLS,[53-59] CHARMM,[60-61] AMBER,[62-63] GROMACS.[52, 64-66] The force field selection for a given problem is important as the accuracy of an MD simulation is limited by the accuracy of the employed force field equations, which vary in the complexity of each term and the number of terms in each equation. For example, Eqn 1.3 can be expanded to represent the form of most first generation force fields:[52]

$$V = \sum_{i}\sum_{i>j}\left\{ f_{ij}\frac{q_i q_j e^2}{r_{ij}} + 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]\right\} + \sum_{bonds} k_r(r-r_0)^2 + \sum_{angles} k_r(\theta-\theta_0)^2 + \\ + \sum_{impropers} k_r(\xi-\xi_0)^2 + \sum_{n=0}^{6} C_n(\cos(\psi))^n \tag{1.4}$$

The first two terms in Eqn. 1.4 describe the non-bonded energy between all atoms *i* and *j* separated by distance r_{ij} in terms of Coulomb (q_i and q_j are atomic partial charges) with proper scaling factor *f* and 6-12 Lennard-Jones interactions (σ and ε are Lennard-Jones radius and well depth constants, respectively). The last four terms are for the energy associated with bond-stretching, angle-bending, molecular planarity (improper dihedrals) and bond-rotation (proper dihedrals). Among them, the first three are modeled as

harmonic potentials with force constants (k_r, k_θ, k_ξ) that describe the energies when internal coordinates (r, θ, ξ) deviate from their set values (r_0, θ_0, ξ_0). The improper dihedral energy term is to describe the energy barriers associated with the out-of-plane motions of planar groups such as aromatic rings. The final terms describe the proper dihedral potentials, the rotations around a single bond, using Ryckaert-Bellemans functions.

1.5.2 mPPE modeling studies and replica exchange molecular dynamics simulations

Several modeling studies[23-24, 27-31, 37] have demonstrated the ability of MD simulations to accurately predict mPPE folding phenomenon, which has aided our understanding of this process. For example, Pande and coworkers[28-31] conducted a series of MD simulation studies to examine the folding mechanism of the 12-mer ester functionalized mPPE in implicit and explicit solvent systems, including water, chloroform, acetonitrile and methanol. Lee and Saven[37] used 18-mer ester mPPE models to investigate the variation of the helical structure in water. Most of the above simulation work used simplified models of the ester functionalized mPPE; specifically, methyl groups replaced the long triethylene monomethyl ether pendants.

Early modeling work by our group involved using MD simulations to predict the stable conformation of mPPEs having both ether and amine functional groups in water solutions.[23-24, 49] A considerable effort was put forth to construct several amine

functionalized mPPE models using the OPLS force field[53-59] and to establish an appropriate MD simulation procedure. Further, our group has demonstrated that replica exchange molecular dynamics (REMD)[67-69] simulations are more effective in simulation time than conventional MD simulations in predicting mPPE folding behaviors.[23-24, 49] The REMD technique is an advanced hybrid method incorporating both molecular dynamics and Monte Carlo (MC) concepts to reduce the computational times necessary to identify minimum energy molecular (polymer) structures (i.e., preferred molecular conformations). In general, it converges much more quickly, compared to a conventional MD simulation, to the thermodynamically favored state of the system.[68-72]

An REMD simulation could be described as series of MD simulations of the same system, a replica, simultaneously and independently conducted at different temperatures. After a specified number of MD steps, the atomic coordinates from a pair of thermally adjacent replicas are potentially exchanged based on an acceptance probability calculated from the Metropolis criterion:[73]

$$P_{ij}=\min\left(1,\exp\left[\left(\frac{1}{k_B T_i}-\frac{1}{k_B T_j}\right)\left(E_j-E_i\right)\right]\right) \qquad (1.5)$$

where *i* and *j* represent two adjacent replicas having potential energies E_i and E_j, respectively; P_{ij} is the acceptance probability for exchange between replicas *i* and *j*; and k_B is the Boltzmann constant.

This exchange scheme allows the system to escape the local energy minima by exchanging replicas at low temperatures with replicas at higher temperatures. Figure 1.7 shows the conceptual illustration of the REMD method.

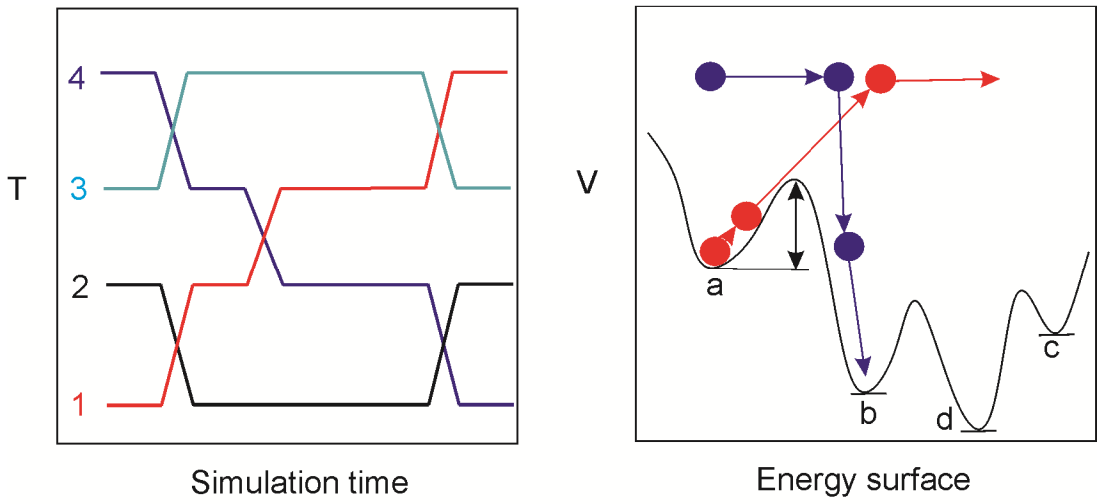


Figure 1.7. Conceptual illustration of the Replica Exchange Molecular Dynamics (REMD) method. The schematic on the left shows the exchange of configuration occuring between thermally adjacent replicas in an REMD simulation. The path of each replica is represented by a different color line. The Metropolis based exchanges allow the simulated system to escape the local energy minima (a, b or c) and to reach the global energy minimum (d) by exchanging the replica at low temperature (red) with the replica at higher temperatures (purple), as shown on the conceptual representation of the energy surface on the right (adapted from Ref. 41).

1.6 Summary

Previously published studies have introduced a range of functionalized mPPE materials, described their synthesis in great detail and examined their propensity to fold into helical conformations. Based on those results, several new applications using the extended and helical conformations of functionalized mPPE were proposed. Further, a number of simulation works exploring this folding concept and its affecting factors have been conducted. These results provide the foundation for the modeling and synthesis studies described in this dissertation, which focus on the *a priori* prediction of polymer structure and novel applications for the functionalized mPPE helical structures.

1.7 References

1. Hill, D. J.; Mio, M. J.; Prince, R. B.; Hughes, T. S.; Moore, J. S., *Chem. Rev.* **2001**, *101* (12), 3893-4011.

2. Nakano, T.; Okamoto, Y., *Chem. Rev.* **2001**, *101* (12), 4013-4038.

3. Nelson, J. C.; Saven, J. G.; Moore, J. S.; Wolynes, P. G., *Science* **1997**, *277* (5333), 1793-1796.

4. Ray, C. R.; Moore, J. S., Supramolecular Organization of Foldable Phenylene Ethynylene Oligomers. In *Poly(Arylene Ethynylene)S: from Synthesis to Application*, **2005**; Vol. *177*, pp 91-149.

5. Brunsveld, L.; Prince, R. B.; Meijer, E. W.; Moore, J. S., *Org. Lett.* **2000**, *2* (11), 1525-1528.

6. Cary, J. M.; Moore, J. S., *Org. Lett.* **2002**, *4* (26), 4663-4666.

7. Gin, M. S.; Moore, J. S., *Org. Lett.* **2000**, *2* (2), 135-138.

8. Hill, D. J.; Moore, J. S., *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (8), 5053-5057.

9. Kelley, R. F.; Rybtchinski, B.; Stone, M. T.; Moore, J. S.; Wasielewski, M. R., *J. Am. Chem. Soc.* **2007**, *129* (14), 4114-4115.

10. Lahiri, S.; Thompson, J. L.; Moore, J. S., *J. Am. Chem. Soc.* **2000**, *122* (46), 11315-11319.

11. Matsuda, K.; Stone, M. T.; Moore, J. S., *J. Am. Chem. Soc.* **2002**, *124* (40), 11836-7.

12. Prest, P. J.; Prince, R. B.; Moore, J. S., *J. Am. Chem. Soc.* **1999**, *121* (25), 5933-5939.

13. Prince, R. B.; Saven, J. G.; Wolynes, P. G.; Moore, J. S., *J. Am. Chem. Soc.* **1999**, *121* (13), 3114-3121.

14. Prince, R. B.; Okada, T.; Moore, J. S., *Angew. Chem. Int. Ed.* **1999**, *38* (1-2), 233-236.

15. Prince, R. B. Phenylene ethynylene foldamers: cooperative conformational transition, twist sense bias, molecular recognition properties and solid-state organization. Ph.D. Dissertation, University of Illinois ar Urbana-Champaign, Urbana-Champaign, IL, U.S.A., **2000**.

16. Prince, R. B.; Barnes, S. A.; Moore, J. S., *J. Am. Chem. Soc.* **2000**, *122* (12), 2758-2762.

17. Prince, R. B.; Brunsveld, L.; Meijer, E. W.; Moore, J. S., *Angew. Chem. Int. Ed.* **2000**, *39* (1), 228-230.

18. Tanatani, A.; Hughes, T. S.; Moore, J. S., *Angew. Chem. Int. Ed.* **2002**, *42* (2), 325-328.

19. Stone, M. T.; Heemstra, J. M.; Moore, J. S., *Acc. Chem. Res.* **2006**, *39* (1), 11-20.

20. Zhao, D. H.; Moore, J. S., *J. Am. Chem. Soc.* **2003**, *125* (52), 16294-16299.

21. Zhao, D. H.; Moore, J. S., *Org. Biomol. Chem.* **2003**, *1* (20), 3471-3491.

22. Zhao, D.; Moore, J. S., *Chem. Commun. (Camb.)* **2003**, (7), 807-18.

23. Adisa, B.; Bruce, D. A., *J. Phys. Chem. B* **2005**, *109* (15), 7548-7556.

24. Adisa, B.; Bruce, D. A., *J. Phys. Chem. B* **2005**, *109* (42), 19952-19959.

25. Arnt, L.; Tew, G. N., *J. Am. Chem. Soc.* **2002**, *124* (26), 7664-7665.

26. Arnt, L.; Tew, G. N., *Macromolecules* **2004**, *37* (4), 1283-1288.

27. Blatchly, R. A.; Tew, G. N., *J. Org. Chem.* **2003**, *68* (23), 8780-5.

28. Elmer, S. P.; Pande, V. S., *J. Phys. Chem. B* **2001**, *105* (2), 482-485.

29. Elmer, S. P.; Pande, V. S., *J. Chem. Phys.* **2004**, *121* (24), 12760-12771.

30. Elmer, S. P.; Park, S.; Pande, V. S., *J. Chem. Phys.* **2005**, *123* (11), 114902(0-14).

31. Elmer, S. P.; Park, S.; Pande, V. S., *J. Chem. Phys.* **2005**, *123* (11), 114903(0-7).

32. Hecht, S.; Khan, A., *Angew. Chem. Int. Ed.* **2003**, *42* (48), 6021-6024.

33. Huang, S. L.; Tour, J. M., *J. Org. Chem.* **1999**, *64* (24), 8898-8906.

34. Huang, Y. Q.; Fan, Q. L.; Liu, X. F.; Fu, N. N.; Huang, W., *Langmuir* **2010**, *26* (24), 19120-8.

35. Khan, A.; Hecht, S., *Chem. Commun.* **2004**, (3), 300-301.

36. Khan, A.; Kaiser, C.; Hecht, S., *Angew. Chem. Int. Ed.* **2006**, *45* (12), 1878-1881.

37. Lee, O. S.; Saven, J. G., *J. Phys. Chem. B* **2004**, *108* (32), 11988-11994.

38. Li, C. J.; Slaven, W. T.; Chen, Y. P.; John, V. T.; Rachakonda, S. H., *Chem. Commun.* **1998**, (13), 1351-1352.

39. Tan, C. Y.; Pinto, M. R.; Kose, M. E.; Ghiviriga, I.; Schanze, K. S., *Adv. Mater.* **2004**, *16* (14), 1208-1212.

40. Zhao, X. Y.; Schanze, K. S., *Langmuir* **2006**, *22* (10), 4856-4862.

41. Lehninger, A. L.; Nelson, D. L.; Cox, M. M., *Lehninger principles of biochemistry*. 5th ed.; W.H. Freeman: New York, **2008**; pp 113-148, 271-301.

42. Carey, F. A.; Sundberg, R. J., *Advanced organic chemistry, Part B: Reactions and Synthesis*. 5th ed.; Springer: New York, **2007**; pp 723-728.

43. Chinchilla, R.; Najera, C., *Chem. Rev.* **2007**, *107* (3), 874-922.

44. Tykwinski, R. R., *Angew. Chem. Int. Ed.* **2003**, *42* (14), 1566-1568.

45. Khan, A. U. H. Phenylene ethynylene foldamers: from synthesis to tubular scaffolding and photoswitchable helices. Ph.D. Dissertation, Freien Universität Berlin, Berlin, Germany, **2005**.

46. Stone, M. T.; Moore, J. S., *Org. Lett.* **2004**, *6* (4), 469-472.

47. Zhao, D.; Moore, J. S., *J. Am. Chem. Soc.* **2003**, *125* (52), 16294-9.

48. Zhao, D.; Moore, J. S., *Org. Biomol. Chem.* **2003**, *1* (20), 3471-91.

49. Adisa, B. Synthesis, characterization and modeling of novel polymers and polymer-templated mesoporous materials. Ph.D. Dissertation, Clemson University, Clemson, SC, U.S.A., **2005**.

50. Allen, M. P.; Tildesley, D. J., *Computer simulation of liquids*. Oxford University Press: Oxford, England, **1987**.

51. Frenkel, D.; Smit, B., *Understanding molecular simulation: from algorithms to applications.* 2nd ed.; Academic Press: San Diego, **2002**.

52. van der Spoel, D.; Lindahl, E.; Hess, B.; van Buuren, A. R.; Apol, E.; Meulenhoff, P. J.; Tieleman, D. P.; Sijbers, A. L. T. M.; Feenstra, K. A.; van Drunen, R.; Berendsen, H. J. C. GROMACS USER MANUAL version 4.5.4 http://www.gromacs.org/documentation (accessed on 08/03/2011).

53. Jorgensen, W. L.; Madura, J. D.; Swenson, C. J., *J. Am. Chem. Soc.* **1984**, *106* (22), 6638-6646.

54. Jorgensen, W. L., *J. Phys. Chem.* **1986**, *90* (7), 1276-1284.

55. Jorgensen, W. L.; Maxwell, D. S.; TiradoRives, J., *J. Am. Chem. Soc.* **1996**, *118* (45), 11225-11236.

56. Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L., *J. Phys. Chem. B* **2001**, *105* (28), 6474-6487.

57. Price, M. L. P.; Ostrovsky, D.; Jorgensen, W. L., *J. Comput. Chem.* **2001**, *22* (13), 1340-1352.

58. Rizzo, R. C.; Jorgensen, W. L., *J. Am. Chem. Soc.* **1999**, *121* (20), 4827-4836.

59. Watkins, E. K.; Jorgensen, W. L., *J. Phys. Chem. A* **2001**, *105* (16), 4118-4125.

60. Kale, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K., *J. Comp. Phys.* **1999**, *151* (1), 283-312.

61. Mackerell, A. D.; Wiorkiewiczkuczera, J.; Karplus, M., *J. Am. Chem. Soc.* **1995**, *117* (48), 11946-11975.

62. Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P., *J. Am. Chem. Soc.* **1984**, *106* (3), 765-784.

63. Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A., *J. Comput. Chem.* **1986**, *7* (2), 230-252.

64. Berendsen, H. J. C.; van der Spoel, D.; Vandrunen, R., *Comput. Phys. Commun.* **1995**, *91* (1-3), 43-56.

65. Lindahl, E.; Hess, B.; van der Spoel, D., *J. Mol. Model.* **2001**, *7* (8), 306-317.

66. van der Spoel, D.; E., L.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C., *J. Comput. Chem.* **2005**, *26*, 1701-1718.

67. Patriksson, A.; van der Spoel, D., *Phys. Chem. Chem. Phys.* **2008**, *10* (15), 2073-2077.

68. Seibert, M. M.; Patriksson, A.; Hess, B.; van der Spoel, D., *J. Mol. Biol.* **2005**, *354* (1), 173-183.

69. Sugita, Y.; Okamoto, Y., *Chem. Phys. Lett.* **1999**, *314* (1-2), 141-151.

70. Beck, D. A. C.; White, G. W. N.; Daggett, V., *J. Struct. Biol.* **2007**, *157* (3), 514-523.

71. Periole, X.; Mark, A. E., *J. Chem. Phys.* **2007**, *126* (1), 014903(0-11).

72. Zhang, W.; Wu, C.; Duan, Y., J. *Chem. Phys.* **2005**, *123* (15), 154105(0-9).

73. Metropolis, N.; Rosenluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E., *J. Chem. Phys.* **1953**, *21*, 1087-1092.

CHAPTER 2

MODELING STUDY OF FUNCTIONALIZED *meta*-POLY(PHENYLENE ETHYNYLENE) FOLDAMERS

2.1 Introduction

The prediction and control of secondary structure in polymers has long been of interest to the scientific community, because the functions and activities of these materials can often be correlated to their structure. This has implications for a wide range of technologies, including self-assembled nanostructures, high-performance engineering materials, and biologically active molecules such as proteins and enzymes. Though considerable progress has been made in understanding secondary structure formation in these systems, the complexity of such molecules has often made it difficult for a reliable, purely predictive framework to be established. Thus, many researchers have sought out simple analogs of these complex systems, so as to understand the fundamental science behind secondary structure formation, and to build a foundation upon which more complex frameworks may be based.

In this effort to develop new biomimetic materials and better understand secondary structure formation, Moore and co-workers[1-10] introduced a new class of polymers named *meta*-poly(phenylene ethynylene)s (mPPEs). Some of these polymers are known to fold into an ordered, helical conformation in solution, and the folding process is known to be

influenced by various factors, including solvent conditions, temperature, etc. Thus, mPPEs have the potential for use in a wide range of applications that require self-assembly (e.g., chemical sensing, biocide coatings, and catalysis), yet they also provide an interesting template for studying the fundamental interactions that govern macromolecular conformations. The folding processes of several mPPE variations have been examined in detail by others,[11-18] and classical molecular simulation methods have proved a useful tool in such studies. Yet, to date, simulation has been used primarily as a means to study known materials, rather than as a predictive tool. In this work, we demonstrate that related simulation methods can be successfully used in predictive studies, allowing a large combinatorial parameter space to be screened on the computer before investing laboratory resources on the synthesis and characterization of new materials.

The primary adjustable features of mPPEs are the functional groups attached to the phenylene ethynylene backbone. Our initial studies examined the structure directing effects of functional groups positioned *meta* to the polymer backbone. As shown in Figure 2.1, these pendant groups may take on a wide range of chemical functionality. In general, the formation of secondary structure in any given mPPE is dependent on the functionality of its side groups (labeled R in Figure 2.1) and the solvent in which the mPPE is placed. For example, the ester-functionalized mPPE has been experimentally proven to fold into a helical conformation in acetonitrile and other solvents, but maintains a random coil conformation in chlorinated solvents.[3, 5]

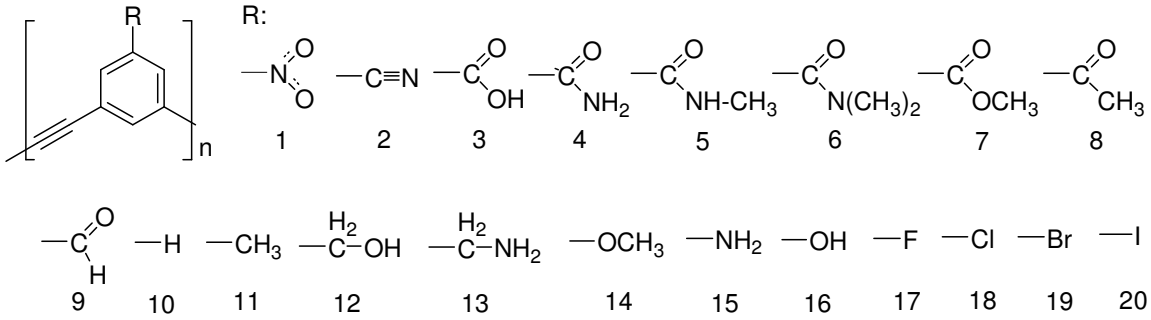


Figure 2.1 Structure of *meta*-poly(phenylene ethynylene) polymers having one functional group on each aromatic ring.

Previous studies of helix formation by mPPEs have identified several structural factors that contribute to folding. First, the series of *meta*-ethynylene linkages in the polymer backbone allows for π/π interactions to occur between aromatic groups that are 1, 7 neighbors, as shown in Figure 2.2. This interaction has a stabilizing effect, which is analogous to the formation of hydrogen bonds in a protein α-helix. This effect may be reinforced or counteracted by other important factors, including interactions between the solvent and solvophilic or solvophobic sites along the polymer, and the presence of electron withdrawing or donating groups attached to the aromatic rings.[2, 4] Continuing with our previous example, the folding preference of the ester-functionalized mPPE in acetonitrile has been attributed to a combination of these three factors: the solvophilic nature of its ester side groups, the solvophobic nature of the polymer backbone, and the electron-withdrawing character of the ester groups on the aromatic rings. Because the compact helix exposes the mPPE side groups to solvent, while shielding the polymer backbone, the helix conformation is quite stable from the perspective of polymer/solvent

interactions. Further, it has been argued that electron-withdrawing groups, such as esters, on the aromatic rings serve to strengthen the π-stacking interactions between polymer segments, adding more stability to the helix.[4]

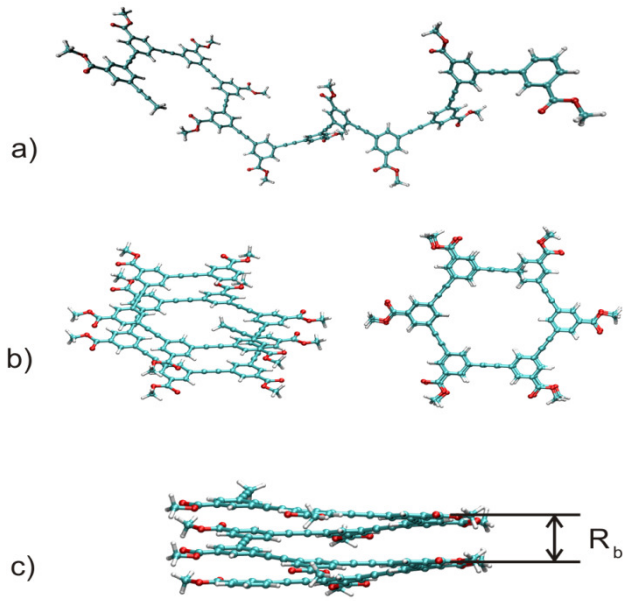


Figure 2.2 Typical conformations of mPPEs. This figure shows the ester mPPE (R = 7) in a) a random extended conformation, and in a helical conformation as viewed b) axially and c) perpendicular to the helical axis. The interlayer or inter-turn spacing equals D_a. Atom colors for carbon, hydrogen and oxygen are teal, white and red, respectively.

The above explanation is certainly adequate to justify the observed folding behavior of an ester-functionalized mPPE, which possesses a folding bias in many solvents.[3, 7] However, such reasoning is not sufficient to establish a reliable set of heuristics for predicting folding behavior *a priori*, since the competing effects of solvophilic/solvophobic

interactions and π-stacking interactions are not easily quantified. For example, the conventional argument asserts that electron donating substituents on the aromatic rings should weaken π-stacking interactions, and therefore, favor the unfolded state; this was offered as a reason why an mPPE with electron-donating groups did not fold in acetonitrile and other polar solvents.[4] However, a water soluble, folding mPPE with electron donating functional groups was successfully synthesized,[19] contrary to these heuristics. Indeed, the effect of electron donating/withdrawing groups itself is not clearly understood, as the π-stacking phenomenon is due to a combination of effects, such as exchange repulsion, induction and dispersion.[20-24] Experimental evidence and high level quantum calculations have even suggested that *all* functionalized aromatic structures have stronger π-stacking interactions than normal benzene rings,[20, 22, 23, 25, 26] indicating that both activating and deactivating functional groups could stabilize mPPE helical structures through their effect on π-stacking interactions.

Further complications arise when considering the effect of solvophobic/solvophilic sites on an mPPE polymer's native conformation. Naturally, these definitions will depend on the characteristics of the solvent, and therefore, the choice of solvent has a direct effect on the formation of secondary structure. For example, studies of the ester mPPE in different solvents indicate that certain solvents - chloroform, dichloromethane and tetrahydrofuran - do not favor the folded conformation, while other solvents - acetonitrile and hexane - favor the adoption of a helical conformation. The stability of extended conformations in chlorinated solvents was explained as the result of strong CH/π interactions between the exposed aromatic rings and solvent molecules. [3, 20, 27] Thus,

when chloroform is used as the solvent, the polymer backbone exhibits some solvophilic character. Yet, the effect of solvent upon folding is not easily separated from the effects of side group functionality. For example, methanol induces folding in the ester functionalized mPPE,[3] while the same solvent promotes extended conformations for another mPPE having electrolytic side groups.[19]

Though some of the science behind the folding driving force in mPPEs remains unsettled, we note that there have been several successful studies[11-18] that examined the folding process of mPPEs by classical simulation methods, such as molecular dynamics (MD). For example, the MD results of Elmer et al. show that the folding time of mPPEs generally vary from 70 ns to 400 ns, and this is in agreement with experimental results.[10, 15] In another study, the same authors conducted MD simulations of ester functionalized mPPEs dissolved in four different explicit solvents - acetonitrile, methanol, chloroform and water - to quantify the time required for secondary structure formation.[17] When chloroform and water were used as a solvent, no folding event was witnessed. Therefore, it was concluded, that chloroform solvated the ester mPPEs in such a way that helical structures were disfavored and that the folding process in these solvents would only occur over long time scales, if at all.[17] This result agrees with experimental data, in which no folding was observed in chloroform.[5]

The aforementioned simulation studies demonstrate a favorable comparison between classical models and experimental evidence. Yet, for the purpose of determining the most favored conformation, the procedure of running a single MD simulation initialized from a

random extended state is by itself unable to provide positive proof of the preferred conformation. For example, failure to observe the folding process in such simulations is insufficient evidence that the folded state is unfavorable, because the time scale for the folding process might simply be longer than the time simulated for a given initial polymer configuration. Thus, while conventional MD simulations are ideal for examining folding rates, they are not well suited for the purpose of this study - determining whether the helical conformation is favored for a given mPPE/solvent pair.

Adisa and Bruce[11-13] used two enhancements to the conventional MD approach in their study of mPPEs, resulting in a method specifically designed to determine the native state of mPPEs in solution. The first enhancement was the use of Replica Exchange Molecular Dynamics (REMD)[28] rather than conventional MD. The REMD technique is a hybrid method that incorporates both Newtonian dynamics and Monte Carlo (MC) concepts, and it is well-suited for use with modern parallel computing architectures. Compared to a conventional MD simulation, REMD converges much more quickly to the thermodynamically favored state of the system,[28-32] allowing the native state to be reached in significantly less simulation time. As a second enhancement, a procedure was used in which two REMD simulations were carried out for each mPPE in solvent: one beginning with an extended initial conformation, and the other with a helical initial conformation. Using results from both starting points, a more decisive conclusion may be reached regarding the native conformation of the system. The authors modeled two mPPE variations with this technique, one with ether and the other with amine side chains, in explicit water. Simulation results indicated that, for both the ether and amine

functionalized mPPEs, helix formation is likely to occur in water.[12] Though the incorporation of replica exchange moves in REMD make the technique unsuitable for studying mPPE folding kinetics and mechanism, our results demonstrate that the method is effective in determining the favored conformation of a given mPPE in solution.

In the present study, a similar protocol was used as a simple, quick test for screening the folding behavior of over 100 different mPPE/solvent combinations. This procedure was first validated against the previously reported experimental data. It was then used to predict the structure of two new variations of mPPEs, which were later synthesized and characterized to confirm the REMD predictions. We then employed this method in a large-scale combinatorial investigation, over a wide range of mPPEs with different functional groups in several solvent conditions. These combined simulation-synthesis efforts clearly show the predictive capabilities of the REMD method and illustrate how these simulations can be used to significantly reduce, if not eliminate, trial and error synthesis efforts focused on the production of ordered organic molecules.

2.2 Computational method

2.2.1 Computer system

All simulations were conducted using the *Palmetto* supercomputer[33] at Clemson University. The Palmetto cluster consists of many different computer systems. The computers used in this study were Dell PE 1950 with 2x Intel Xeon E5345 Quad Core

processors at 2.33 GHz, 4MB L2 Cache, 12GB RAM and 80GB of local storage (with 120TB of network storage). The computers and network are linked by 10G-SW32LC-16M Myrinet linecards.

REMD simulations in this study employed from 24 to 64 CPUs, and for each simulation it took 32 – 72 h of computing resources to complete 10 ns of simulation time using the Gromacs computational engine (1 CPU for each replica).

2.2.2 Simulation procedure

All MD and REMD simulations were conducted using the freeware program Gromacs, version 3.3.1 (the fastest MD software currently available).[34-40] The initial mPPE structures were generated using Materials Studio 4.4.[41] An example of the process used to generate an mPPE structure using Materials Studio is presented in Appendix A. The procedure for preparing and minimizing the solvated box was similar to that reported by Adisa and Bruce.[11-13] All bond lengths in the polymers were held constant using the LINCS algorithm.[36] The solvated boxes were equilibrated using isothermal-isobaric (*NPT*) molecular dynamics for 200 ps, with Berendsen[42] temperature coupling (T = 300 K, τ_T = 0.1 ps) and Parrinello-Rahman[43] pressure coupling (P = 1 bar, τ_P = 1 ps). The resulting equilibrated systems were then used as the initial structures for REMD simulations at constant volume (*NVT*).

To determine the preferred structure of each mPPE, two REMD simulations were performed: one beginning with the polymer in a random extended conformation, and the other beginning with a helical conformation. For the first simulation, the conformation of the molecule is not uniquely defined, since the extended state is essentially random. Yet, preliminary tests consisting of REMD simulations initialized with ten unique mPPE conformations showed little influence of the initial polymer conformation on the observed folding behavior of the ester mPPE. Therefore, for all remaining mPPEs, only one random extended conformation was used. Thus, we found that two simulations - one beginning with a random structure and the other beginning with a helical structure - are sufficient to determine whether a given mPPE is likely to take on a helical or random extended structure in a studied solvent condition.

2.2.3 REMD simulation parameters

REMD simulations were conducted using a temperature range of 300 K to 600 K. The temperature distribution and the number of replicas were selected based on the procedure described by Adisa and Bruce,[11, 12] such that the exchange probabilities in all simulations were approximately 15% to 20%.[12, 30, 32, 44] A time step set of 2 fs was used, with replica exchange moves attempted every 0.5 ps (250 steps). The length of each REMD simulation was kept at 10 ns. The other simulation parameters used here were identical to those reported earlier.[12] Twin range cut-offs were used for van der Waals and Coulomb interactions, each with a cut-off radius of 1.2 nm; neighbor lists were updated every 10

steps; atom trajectories and energy terms were saved every 5000 steps. Additional simulation details are presented in Appendix B.

Though a range of temperatures are examined in REMD simulations, only the trajectories at 300 K were used in our analysis, representing the mPPEs structures at ambient temperature. Graphical representations of mPPE structures were obtained with Visual Molecular Dynamics.[45] To evaluate the level of order in mPPE conformations during REMD simulations, the following parameters were calculated: the radius of gyration (R_g, nm), the solvent accessible surface area(*SASA*, nm^2),[46, 47] the Lennard-Jones interaction potential between mPPE atoms ($V_{LJ,P\text{-}P}$, kJ/mol), the Lennard-Jones interaction potential between mPPE atoms and the solvent molecules ($V_{LJ,P\text{-}S}$, kJ/mol), and the distance between two overlapping aromatic rings (D_a, nm) as defined in Figure 2.2c. Because of the similarity in time evolution of these five parameters, they were reported only with the first two results. Subsequently, R_g was selected as the preferred order parameter, and calculated using the following equation:[40]

$$R_g = \left(\frac{\sum_i m_i r_i^2}{m_i} \right)^{1/2} \tag{2.1}$$

where m_i is atomic mass and r_i is the position of atom i with respect to the center of mass of the molecule.

Other structural parameters to monitor REMD simulations, such as the consecutive number of π-stacking pairs and the consecutive number of *cisoid* conformations, are introduced in Appendix C and used in subsequent mPPE simulation studies.

2.2.4 Functionalized mPPE models

The atom types, bonded and non-bonded parameters for the mPPE models were taken directly from the Optimized Potential for Liquid Simulations (OPLS),[48-54] as distributed with Gromacs 3.3.1, with some modifications.[11, 13] An example of these mPPE models is presented in Appendix D. A polymer chain length of 12 monomers (12 aromatic rings) was selected, based on the length required for a stable helical structure in experimental studies.[5, 9] To simplify the model, the long solvophilic pendants used in experimental studies[5] were excluded, based on evidence published in previous studies that their role in mPPE folding behavior is likely negligible.[15, 17] In experimental studies, the lengths of those solvophilic tails were designed to facilitate solvation of the mPPEs, and this consideration is less important for molecular simulations because a single polymer chain may be placed in a simulation environment (or unit cell) where it is surrounded by solvent molecules at conditions that mimic infinite dilution.

The mPPE models were built according to the procedure described earlier.[11, 12] The extended structures were built by randomly assigning torsion angles (dibenzyl ethynylene torsion angles), and the helical structures were built by manually adjusting dihedral

angles in the extended structure until a helical conformation was obtained. In constructing the helices, the distances between two overlapping aromatic rings were set at 0.35 nm, following the optimum distance of π/π interactions in benzene and its functionalized derivatives.[20, 22, 55] In total, twenty mPPEs were built for the combinatorial portion of this study, using the functional group variations shown in Figure 2.1.

2.2.5 Solvent models

Eight solvents were considered in this study. These were selected to represent a range of polarity, from highly polar to nonpolar, as described by Reichardt.[56] Table 2.1 lists these solvents and their isothermal compressibilities and dielectric constants, as reported in literature.[57-59] Models for these molecules were also taken from the OPLS all atom (AA) and united atom (UA) force fields.[48-54] An example of these solvent models is presented in Appendix C. In all *NPT* simulations, the listed experimental compressibilities were used for defining the pressure coupling time constant τ_P, with arithmetic averages used for bicomponent mixtures. All solvents, except chloroform, were represented by united atom models that effectively combined aliphatic hydrogens and their respective carbon atom into a single pseudo atom. In this study, the ester mPPE (R = 7) was simulated in all eight solvents, as well as in three mixtures of acetonitrile and chloroform, as a means of validating the modeling approach against existing experimental data.[5, 6] Other mPPEs were simulated in five solvents: chloroform, acetonitrile, acetone, diethyl ether and hexane. We note that the UV absorbance of acetone does not allow an effective

spectroscopic study of folding; thus, the simulation results presented in this work provide a means to investigate the folding behavior of mPPEs in this type of solvent, a task which would otherwise be difficult to accomplish.

Table 2.1 Solvents used in modeling efforts examining the folding behavior of mPPEs.

Solvent	Isothermal compressibility[a] ($x10^5$ bar^{-1})	Dielectric constant[c]
Acetonitrile	10.7	36.64
Chloroform[b]	11.28	4.8069
Diethyl ether	20.90	4.2666
Methanol	12.93	33.0
Ethanol	11.91	25.3
Hexane	17.09	1.8865
Carbon tetrachloride[b]	11.28	2.2379
Acetone	12.23	21.01

[a] From Lide and Kehiaian[57] and Torres et al.[58]
[b] All-atom model (others are united-atom models).
[c] From Wohlfarth.[59]

2.3 Results and discussion

2.3.1 Categorization of REMD simulation results

The REMD simulation results for each mPPE were found to fall under one of four classifications. As indicated in Figure 2.3, there are several possible outcomes for each pair of simulations. If the favored state of the mPPE is unfolded, the simulation initialized

from the folded state will show an increase in R_g upon unfolding, while R_g will remain relatively constant in the simulation initialized from the unfolded state. We refer to this scenario as Group 1. Alternatively, the REMD simulation with an unfolded initial state may show a decrease in R_g and occasionally form a helix without maintaining the structure for any long period of time. This indicates the helix is slightly stable but not the preferred state, and we refer to this scenario as Group 2. In a third scenario, Group 3, the simulation beginning from the helical conformation remains helical during the entire simulation, while the one initialized with the unfolded structure never forms a stable helix (though helical conformations may be sampled briefly during the simulation). The distinction between Groups 2 and 3 is that the helix is much more stable in Group 3 than in Group 2. Further, simulation results in Group 3 are sufficient to indicate folding, though formally no conclusion may be drawn between fully folding and partially folding. As a fourth possibility, the unfolded polymer may progress to a stable helix during the simulation, while the simulation starting from the helix remains unchanged. This scenario clearly indicates the helix is the preferred conformation, and is labeled Group 4. Thus, in this classification system, each mPPE/solvent pair is given a number from 1 to 4, with 1 indicating the least likelihood of the mPPE polymer folding into a helical conformation and 4 the greatest. Representative time evolutions of R_g for these four groups are shown in Figure 2.4.

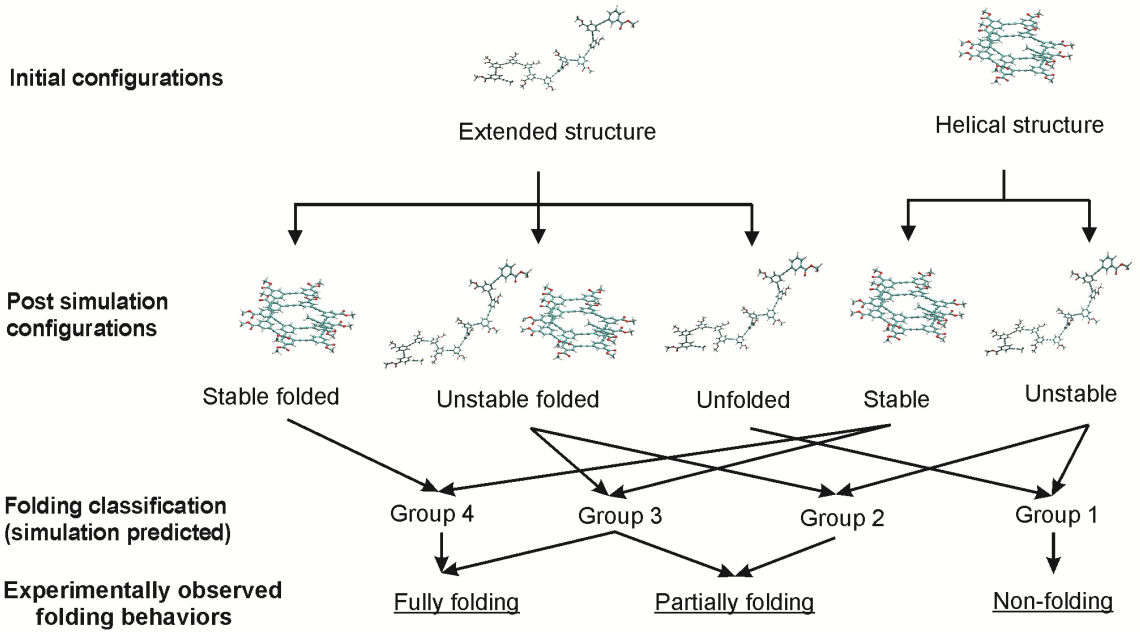


Figure 2.3 Classification of REMD simulation outcomes for mPPE polymers, shown with the corresponding folding behaviors most likely to be observed in experiments. For each mPPE system, two REMD simulations are conducted, one begins with an extended structure and the other begins with a helical structure. Atom colors for carbon, hydrogen and oxygen are teal, white and red, respectively.

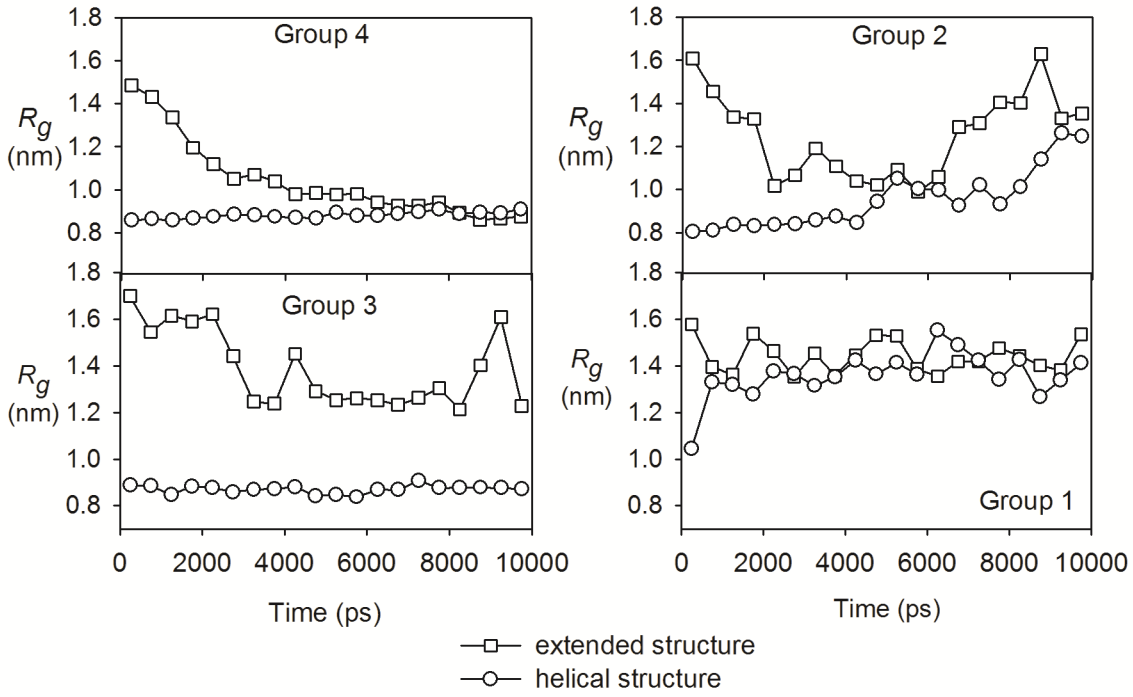


Figure 2.4 Radius of gyration (R_g) data for examples of the different classifications of mPPE simulation results: Group 4 (fully folding, stable helical conformation, R_g = 0.8 to 1.0 nm), Group 3 (partially folding, helical conformation is observed for extended periods, but may not be the only low energy conformation), Group 2 (partially folding, helical conformation is one of many observed conformations), and Group 1 (non-folding, stable unfolded state, R_g > 1.2 nm). Squares indicate simulations beginning from a random coil, and circles indicate simulations beginning from a helical conformation. Each plotted point is a time average over 500 ps, placed at the center of the corresponding time interval.

Note that Group 3, as defined in Figure 2.3, is technically inconclusive to determine the favored structure of the relevant mPPE. If R_g remains relatively constant in both types of REMD simulations, even if several folded states were sampled in the simulation initialized with an extended mPPE structure, this is most likely an indication that the simulations have simply not run long enough to observe the progression to the preferred state. Still, it is clear that the helical conformation is significantly more stable for species ascribed to Group 3 as compared to those in Group 2, and therefore, we may assert that for an mPPE/solvent system in Group 3, folding is likely but not certain.

2.3.2 The folding behavior of the ester functionalized mPPE (R = 7)

The most extensively studied mPPE that exhibits a tendency to fold into a helical conformation is the ester-functionalized mPPE. Its folding behavior has been experimentally characterized in many different solvents, including acetonitrile and chloroform.[3, 5, 6] In this section, our simulation protocol was verified against these experimental results, using an analogous structure (R = 7, see Figure 2.1). The results serve as justification that the simulation procedure used in this study is adequate to evaluate mPPE folding behaviors.

Simulation results for the ester mPPE in acetonitrile and in chloroform fell under Groups 4 and 1, respectively, in agreement with experimental observations - the ester mPPE formed a helical conformation in acetonitrile and a random extended structure in

chloroform. Simulation results are represented by the visual snapshots of the trajectories in Figure 2.5. Further results are given in Figure 2.6, which shows the time evolutions of the R_g, *SASA*, the Lennard-Jones interaction energies and the π-stacking distance D_a.

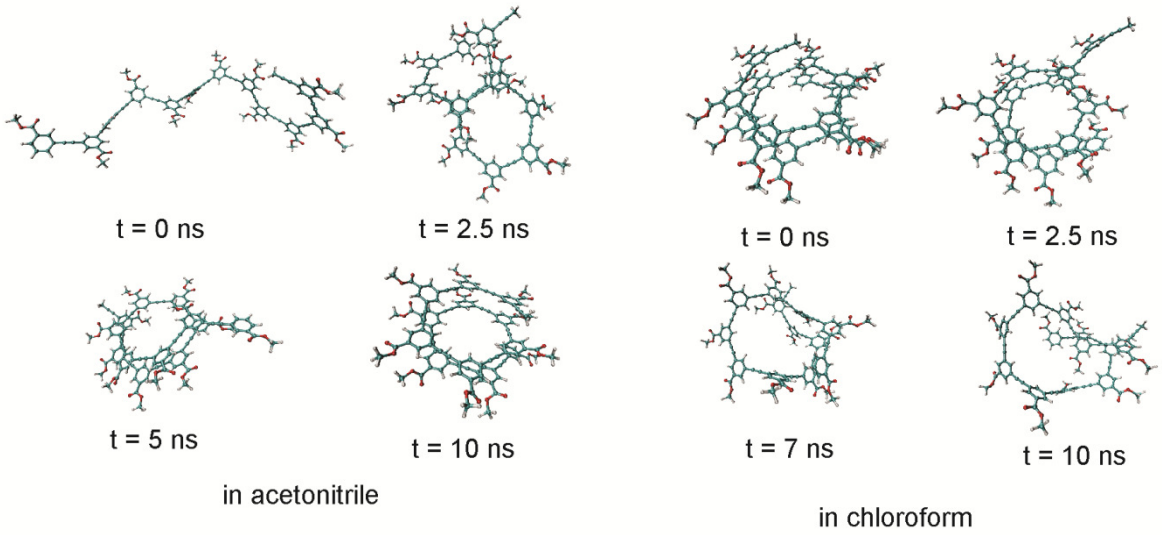


Figure 2.5 Representative conformations at 300 K during REMD simulation of the ester-functionalized mPPE (R = 7), observed at various simulation times. Conformations in acetonitrile were taken from a simulation initialized from a random coil, while the conformations in chloroform were taken from a simulation initialized from a helical conformation. Atom colors for carbon, hydrogen and oxygen are teal, white and red, respectively.

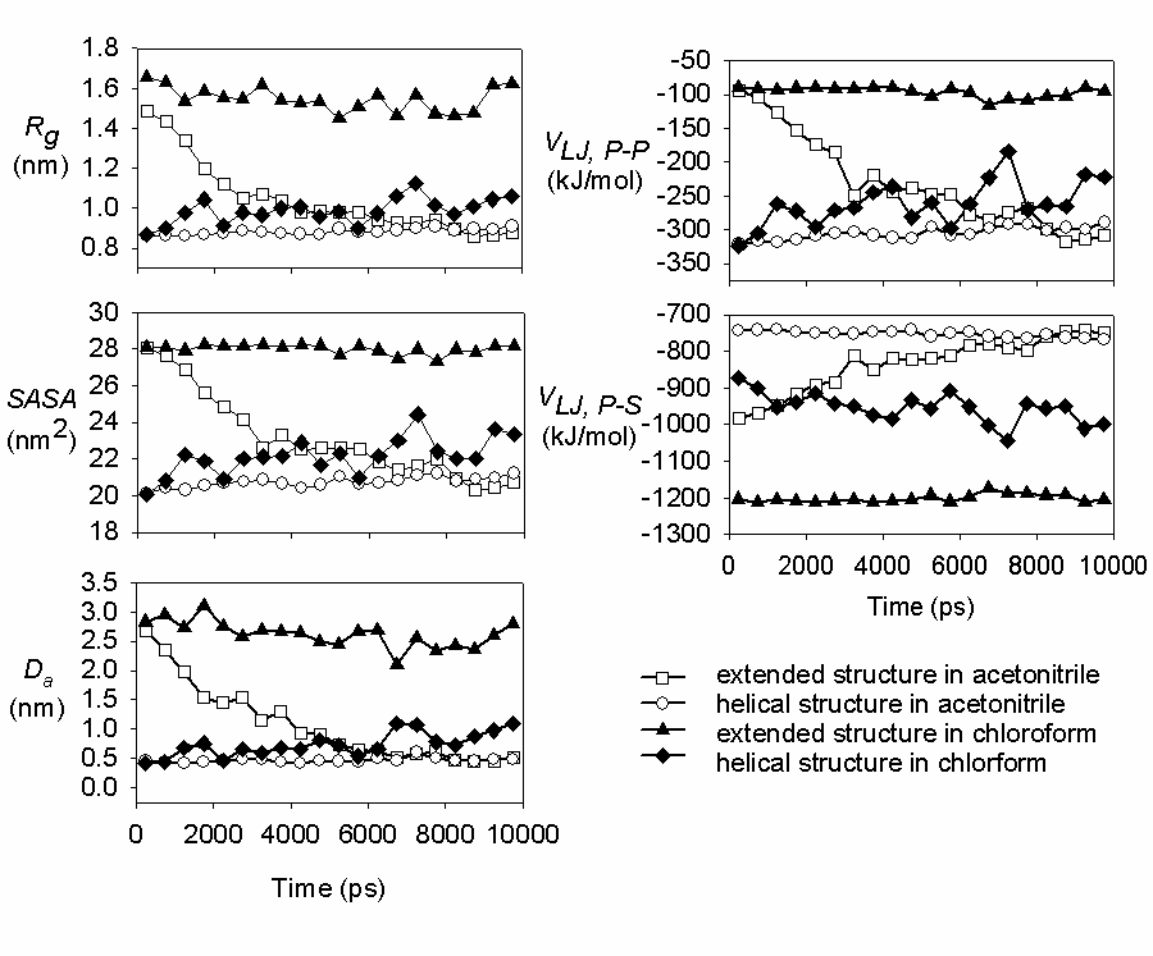


Figure 2.6 Evolution of folding indicators during REMD simulations of the ester functionalized mPPE (R = 7) in acetonitrile (black series) and chloroform (grey); R_g: radius of gyration; *SASA*: solvent-accessible surface area; D_a: phenyl ring separation distance; $V_{LJ,P\text{-}P}$: polymer/polymer Lennard-Jones energy; $V_{LJ,P\text{-}S}$: polymer/solvent Lennard-Jones energy. Each plotted data point is a time average over 500 ps, placed at the center of the corresponding time interval.

The REMD simulation beginning with the extended structure of the ester-functionalized mPPE in acetonitrile indicated that the oligomer formed a stable and well defined helical

conformation. The steadily decreasing R_g and *SASA* indicate the transformation from the extended structure to a more compact conformation, which is also supported by the change in the Lennard-Jones interactions. The intramolecular Lennard-Jones energy of the polymer, $V_{LJ,P\text{-}P}$, became increasingly negative, reflecting an increase in the number of close contacts between segments of the polymer chain. Meanwhile, the polymer/solvent Lennard-Jones energy, $V_{LJ,P\text{-}S}$, became less negative, indicating fewer interactions between the polymer's atoms and solvent molecules. All five parameters became stable after 5 ns, indicating that the compact helix is likely the favored structure. This includes the order parameter D_a, which became nearly constant after 5 ns, confirming the formation of an ordered, defined structure. When the simulation was initialized with a helical mPPE conformation, we observed a stable structure, with only small deviations in all five monitored values over the entire 10 ns simulation. Thus, both simulations indicate that the helix is the thermodynamically favored conformation in acetonitrile.

The above results are sufficient to place the ester mPPE/acetonitrile system in Group 4, according to the classification system defined in Figure 2.3. We note that, in the REMD simulation beginning with an extended structure, the helical conformation was observed as early as 3 ns, becoming stable after 5 ns of simulation time. Compared to the 200 to 400 ns folding time required with traditional MD simulations,[15, 17] this is a significant simulation speedup - even after factoring in the number of extra processors required for the higher temperature replicas in REMD. Because of the accelerated sampling achieved by replica exchange moves, folding times quoted for the REMD results do not correspond to real folding times predicted via conventional MD or observed experimentally;

however, there is evidence that the mPPEs typically do take hundreds of nanoseconds to fold.[10, 17] Yet, in quoting these numbers, we simply wish to demonstrate that REMD requires significantly less computational resources than MD to arrive at the thermodynamically favored state.

Further observations confirm the efficiency of the REMD method in terms of conformational sampling. Before reaching the final helical conformation in acetonitrile, our simulations of the ester mPPE sampled multiple intermediate states similar to those that occurred during the classical MD simulations reported in earlier studies.[15, 17] These intermediate states do not represent the actual pathway or time scale necessary to form a helical conformation from an extended structure, as would be obtained from conventional MD,[15, 17] yet their occurrence implies that the REMD method achieves efficient sampling of the canonical phase space distribution, which is necessary for convergence towards the most stable structure.

In contrast to the steady collapse of the extended structure of the ester mPPE in acetonitrile, our simulations show the extended structure to be favored when the mPPE is in chloroform. The monitored parameters shown in Figure 2.5 indicate the extended mPPE structure did not become more compact in this solvent, suggesting the random extended conformation is favored. This observation was further supported by results from the simulation beginning with the helical structure, in which the helix became unstable and unfolded (see Figure 2.6). An unfolded structure was observed shortly after the beginning of the simulation, at 1.5 ns. From that point forward, the mPPE alternated

between helical and random extended structures with increasing frequency over time, becoming a fully random structure at the end of the simulation. Thus, we conclude that the favored state is the unfolded random extended structure, and categorize this system in Group 1. These results clearly show the agreement between these REMD simulation results and previous experimental results[5] regarding the folding behavior of ester functionalized mPPEs in pure acetonitrile and pure chloroform.

The results in Figure 2.6 show that our selected conformational indicators are closely correlated with each other, and also with the visual depictions of conformations in Figure 2.5. Thus, although all five parameters were examined in each of the following REMD simulations, we have selected the radius of gyration (R_g), as the preferred indicator for all remaining plots.

Simulation results for the ester-functionalized mPPE (R=7) using mixtures of acetonitrile and chloroform (25%, 50%, and 75% acetonitrile by volume), demonstrated a qualitative correlation between the stability of the mPPE helical structure and the concentration of acetonitrile (Figure 2.7). It can be seen that, upon increasing the chloroform concentration, the amount of simulation time required for the extended mPPE structure to become folded increases, with mPPEs in solvent mixtures containing greater than 50% chloroform never reaching the folded state. At 75% chloroform, the simulation of the helical structure begins to show signs of instability, with the extended structure clearly favored in pure chloroform. These observations are in agreement with the experimental results reported earlier.[5, 6]

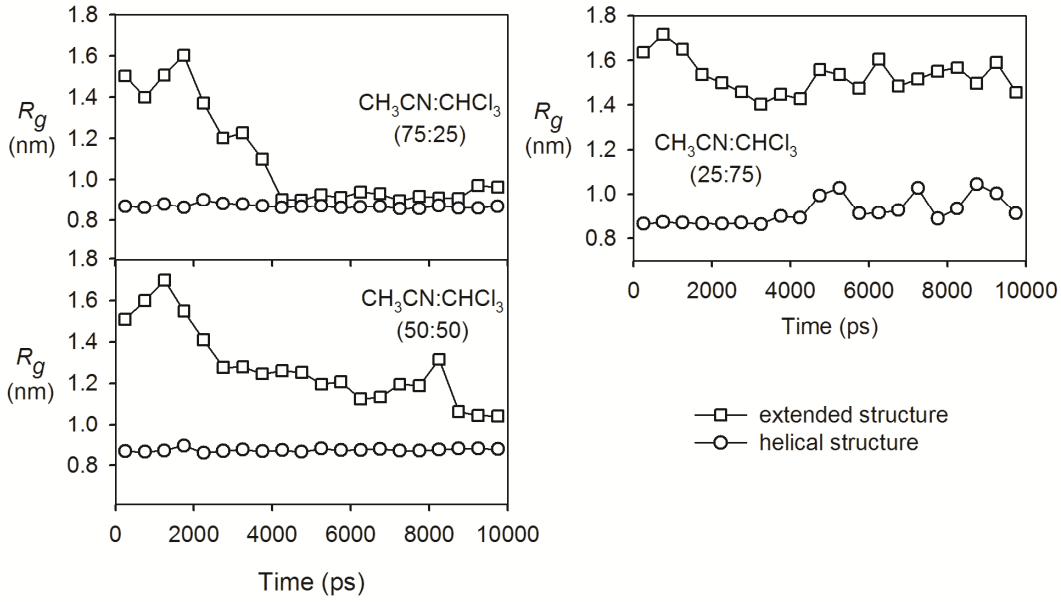


Figure 2.7 Time evolution of the radius of gyration (R_g) of the ester-functionalized mPPE (R = 7) at 300 K, observed during REMD simulations of varying solvent composition (volume ratios listed in insets). Results in pure acetonitrile and in pure chloroform are given in Figure 2.6. For each solvent condition, two REMD simulations were performed: one initialized with the helical conformation (circles) and another initialized with a random configuration (squares). Each plotted point is a time average over 500 ps, placed at the center of the corresponding time interval.

Simulation results also showed that the ester mPPE folded in other solvents, including methanol and carbon tetrachloride (Figure 2.8). This folding bias is consistent with published experimental results.[3] The ester mPPE also folded well in acetone, providing the first evidence that the ester mPPE will form a helical secondary structure in this solvent. Simulation results also showed that ester mPPEs are likely to fold in diethyl

ether, ethanol and hexane, although experimental studies[3] indicated the polymer does not dissolve appreciably in these solvents. Because the ester mPPE simulated in this study has short ester side chains, and the simulation systems consisted of only one mPPE molecule, these results provide evidence suggesting that the long-tailed solvophilic pendants do not play a significant role in the helix formation process. Rather, their function is primarily to increase the solubility of the ester mPPE, as suggested elsewhere.[5, 60] These results also provide justification for our choice to exclude the long solvophilic pendants in the ester mPPE structures in this study.

Although the solvents were selected to investigate the effect of their polarity[56] on the folding behavior of ester mPPEs, no clear difference was observed. The helical structures of the ester mPPE were stable in all selected solvents, except chloroform. The folding behavior of mPPE extended structures appeared to follow a similar overall change from initial structures into the final coiled conformations, though there were minor qualitative differences, which could be attributed to the selection of initial configurations and the number of atoms in each simulation. Similar to acetonitrile, methanol seemed to be a better solvent for inducing folding in the ester mPPE than carbon tetrachloride, as suggested by Hill and Moore.[3]

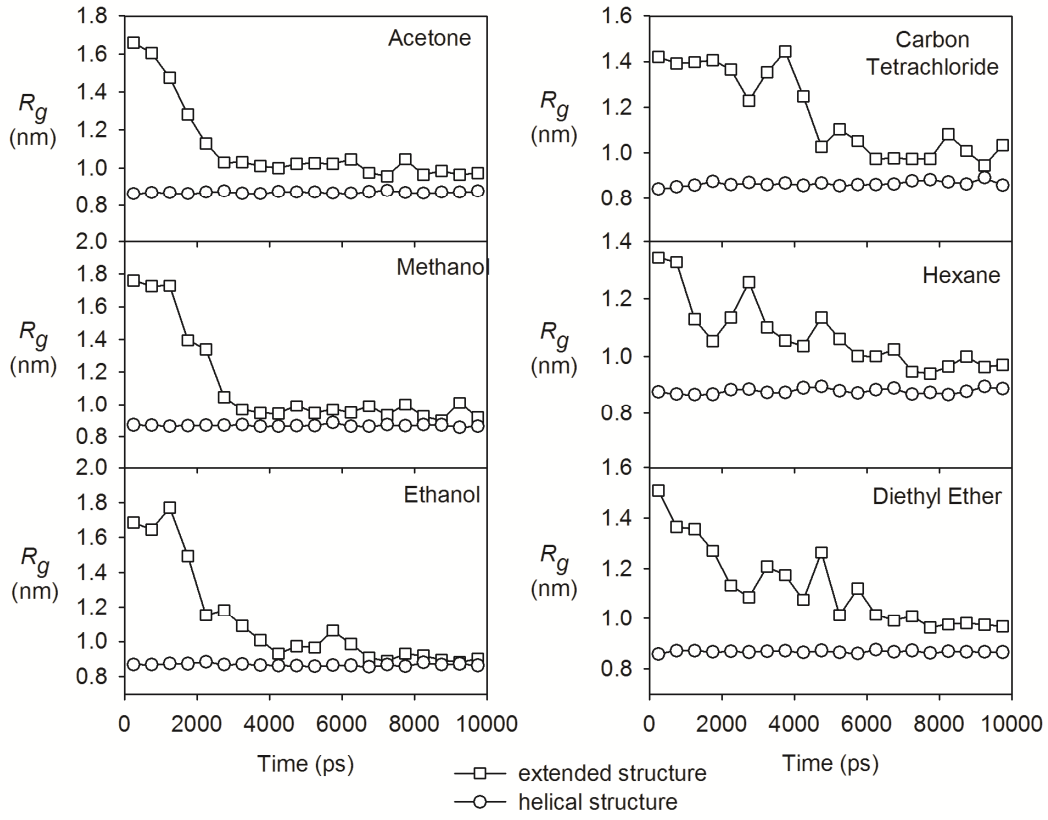


Figure 2.8 Time evolution of the radius of gyration for the ester-functionalized mPPE (R = 7) in various solvents. Each plotted point is a time average over 500 ps, placed at the center of the corresponding time interval.

The helical structures in our simulations have overlapping aromatic rings that are separated by an average distance (D_a) of 0.45 nm, in the range of the optimal π-stacking distance of aromatics rings. This is in reasonable agreement with other high level quantum simulations[20, 22, 55] and experimental data,[61, 62] demonstrating that structurally the helical conformations in our simulations closely resemble those of synthesized mPPEs. This also provides further evidence that the weak π-stacking interactions are well

described by the semi-empirical OPLS force field,[50, 63] justifying our use of the OPLS force field to simulate mPPE systems.

2.4 Combinatorial simulation study

In addition to the ester functionalized mPPE (R = 7), nineteen other mPPE variations (see Figure 2.1) were simulated in chloroform, acetonitrile, acetone, diethyl ether and hexane. The REMD simulation results, including those of the ester mPPE, are compiled in Table 2.2.

The results in Table 2.2 show an excellent agreement between the simulated folding behaviors of mPPEs and the available experimental data for similarly functionalized mPPEs. Experimental results have been previously published for eight of the one hundred entries in Table 2.2, for the mPPEs with R = 7, 12, 13, and 14 in acetonitrile and chloroform, and additionally for R = 7 in hexane. Of these data, referenced in the footnotes of Table 2.2 and Table 2.3, two mPPE/solvent combinations resulted in helical secondary structure formation, one resulted in partial folding behavior, and six indicated no secondary structure. All of these observations are in accord with our simulations results.

Table 2.2 Folding behavior categorization of nine functionalized mPPEs (with functional groups R from 1 to 9, see Figure 2.1) in different solvents, as predicted by REMD simulations.[a]

Functional groups of simulated mPPEs[b]		Folding behaviors observed in different solvents				
		Chloro-form	Aceto-nitrile	Acetone[b]	Diethyl ether	Hexane
$-NO_2$	(R=1)	1	2	2	2	2
–CN	(R=2)	1	2	2	2	1
–COOH[g]	(R=3)	1	3	3	3	2
$-CONH_2$	(R=4)	3	3	3	4	4
$-CONHCH_3$	(R=5)	1	3	3	3	4
$-CON(CH_3)_2$	(R=6)	1	2	3	4	4
$-COOCH_3$	(R=7)	1(1)[d]	4(4)[d]	4	4	4(4)[e]
–CHO	(R=8)	1	2	3	3	3
$-COCH_3$	(R=9)	1	2	3	3	3

[a] Simulation categorization: 1 (non-folding), 2 and 3 (partially folding) and 4 (fully folding). Where possible, experimental observations of mPPEs having similar functional groups are reported in parenthesis.

[b] The functional groups are placed in decreasing order of deactivating effect on the electrophilic aromatic substitution reactions, from $-NO_2$ (R = 1) to $-COCH_3$ (R = 9); in increasing order of activating effect from –H (R = 10) to –OH (R = 16); the halide functional groups (R = 17-20) are placed separately because they have weak deactivating effects but still direct substituents to *ortho* and *para* positions in electrophilic aromatic substitution reactions on the aromatic ring.[64]

[c] No experimental result is listed for mPPEs in acetone, because its UV absorbance prevents measurement of UV and fluorescence spectra.

[d] Based on the ester mPPE with long ether tails studied by Nelson et al.[5]

[e] The ester mPPE with long alkyl tails was shown to fold in hexane by Brunsveld et al.[65]

Table 2.3 Folding behavior categorization of eleven functionalized mPPEs (with functional groups R from 10 to 20, see Figure 2.1) in different solvents, as predicted by REMD simulations.[a]

Functional groups of simulated mPPEs[b]		Folding behaviors observed in different solvents				
		Chloro-form	Aceto-nitrile	Acetone[b]	Diethyl ether	Hexane
–H	(R=10)	1	1	1	1	1
$–CH_3$	(R=11)	1	1	1	1	1
$–CH_2OH$	(R=12)	1(1)[f]	2(2)[f]	2	3	4
$–CH_2NH_2$	(R=13)	1(1)[g]	1	2	2	2
$–OCH_3$	(R=14)	1(1)[h]	1(1)[h]	1	2	2
$–NH_2$	(R=15)	1	1	1	1	2
–OH	(R=16)	1	1	1	1	2
–F	(R=17)	1	1	1	1	1
–Cl	(R=18)	1	1	1	3	3
–Br	(R=19)	1	3	3	3	3
–I	(R=20)	1	3	4	4	2

[a] Simulation categorization: 1 (non-folding), 2 and 3 (partially folding) and 4 (fully folding). Where possible, experimental observations of mPPEs having similar functional groups are reported in parenthesis.

[b] The functional groups are placed in decreasing order of deactivating effect on the electrophilic aromatic substitution reactions, from $–NO_2$ (R = 1) to $–COCH_3$ (R = 9); in increasing order of activating effect from –H (R = 10) to –OH (R = 16); the halide functional groups (R = 17-20) are placed separately because they have weak deactivating effects but still direct substituents to *ortho* and *para* positions in electrophilic aromatic substitution reactions on the aromatic ring.[64]

[c] No experimental result is listed for mPPEs in acetone, because its UV absorbance prevents measurement of UV and fluorescence spectra.

[f] An mPPE with equivalent functional groups showed similar behaviors in the study by Lahiri et al.[4]

[g] A similar mPPE with amine and ether functional groups was studied by Arnt and Tew.[66]

[h] The acid functionalized mPPE was reported by Li et al.[67]

Also reflected in the simulation data is the general difficulty in separating the competing effects of solvophobic/solvophilic interactions and π-stacking strength. While some general trends were evident with respect to these factors, as will be discussed shortly, notable exceptions appeared in every case. The many exceptions are likely attributable to other factors that are less intuitive but equally important, such as site-specific interactions and entropic contributions to the solvation free energy. Such factors are incorporated naturally into the REMD simulations, and are indeed one of the major advantages to using molecular simulation in lieu of heuristics.

When examining trends with respect to solvophobic/solvophilic interactions, a common point of interest is the polarity of the solvent and solute. Generally, polar solvents tend to have favorable interactions with polar solutes, while nonpolar solvents form more favorable solutions with nonpolar solutes. In the context of mPPEs, this argument suggests that helix formation should be stabilized for mPPEs where the pendant groups have similar polarity to the solvent, since in the helical conformation these groups form an outer layer that shields the polymer backbone from the solvent. Likewise, one would expect helix formation to be destabilized by weakly polar and nonpolar solvents, as these should have more favorable interactions with the polymer backbone. Yet, although twenty different functional groups were selected to represent a broad range of polarity, no consistent trend was observed when attempting to correlate the polarity of the solvents and functional groups with the folding behaviors of their respective mPPEs.

For two of the strongly polar mPPE functional groups, nitrile and carboxylic acid, REMD simulations predict that secondary structure formation is inhibited by hexane. This is consistent with the conventional heuristic regarding polarization, outlined above, which indicates that folding should not be likely in such a case. However, the majority of our results indicate that mPPEs are generally *more* likely to fold in nonpolar solvents, such as hexane and diethyl ether, than in polar solvents like acetonitrile or acetone. This exception even applies to some mPPEs with functional groups of high to moderate polarity, such as the amines, amides, and alcohols. Many of these same polar functionalized mPPEs were found to favor extended conformations in the more polar solvents. This is rather unexpected, since the helical structures should be stabilized by their presumably solvophilic coat of polar functional groups, as proposed in earlier studies.[4, 5, 7]

Polarity and dielectric considerations also do not suffice to explain why most mPPEs did not exhibit helical secondary structures in chloroform. Among the twenty functionalized mPPEs in this study, nineteen were clearly unable to form a stable helical structure in chloroform, which is in agreement with experimental data that chlorinated solvents do not favor helical mPPE conformations. Yet, many of these exhibited at least partial folding in acetonitrile, acetone, diethyl ether, and hexane, representing solvents of both greater and lesser polarity than chloroform.[56] This confirms previous assertions that strong CH/π interactions between chloroform and mPPE aromatic groups uniquely stabilize disordered mPPE conformations. [3, 20, 27] Our data suggests only one exception to this behavior, the amide functionalized mPPE (R = 4), which showed partial folding

behavior (Group 3) in chloroform. This anomalous behavior likely results from strong inter-turn hydrogen bonds that form between the amide functional groups of overlapping residues, as observed during simulations of the amide functionalized polymer in its helical conformation. For every other solvent studied, a wide range of folding behaviors were observed over the set of simulated mPPEs.

The simulation results in Table 2.2 and Table 2.3 indicate that, while secondary structure formation is sensitive to solvent conditions, many mPPEs exhibit fairly consistent folding behaviors in a range of solvent environments. For example, out of the twenty mPPEs studied, the ester functionalized mPPE (R = 7) showed the most consistent bias towards the folded state, falling into Group 4 in all solvents but chloroform. Three other mPPEs (R = 10, 11, and 17) showed an equally consistent bias, but towards the unfolded state. These results suggest that each mPPE possesses some innate tendency to fold or remain disordered, independent of the solvent type. One may conclude that intramolecular interactions within the polymers have a slightly stronger effect on secondary structure formation than do the intermolecular solvent-polymer interactions, though clearly most mPPEs show variation with respect to both.

A key intramolecular interaction for helix formation in mPPEs is the π-stacking between aromatic rings, and it is well known that this interaction is affected by the electron withdrawing or donating character of the aromatic substituents.[2, 4] Although the OPLS force field employs a classical point-charge model, it has been shown to reproduce π-stacking energies in substituted aromatics quite well, in some cases better than high level

quantum mechanics computations.[50, 63] Simulation results did indicate a reasonable correlation between the deactivating quality of functional groups and their respective mPPE folding behaviors. All mPPEs with deactivating functional groups, to a certain degree, showed a tendency to form more compact conformations. On the other hand, only a few of the mPPEs having activating functional groups were able to form helical structures in any of the solvents. Nearly all helix forming mPPEs (Group 4) in Table 2.2 and Table 2.3 have deactivating functional groups, with the only exception being the alcohol functionalized mPPE (R = 12), which was predicted to fold only in hexane.

It is rather curious that, despite the general trend with respect to electron withdrawing/donating substituents, the two most strongly deactivating functional groups (R = 1 and 2) exhibited only partial folding in most simulations, or remained unfolded. According to conventional thought, the presence of these groups should enhance the π-stacking interactions, further increasing the stability of the helical structures. These unexpected results bear more similarity to the simulation results of Blatchly and Tew,[68] in which the stability of π-stacking interactions in *ortho*-PPEs were found to be stronger between rings with activating substituents than between rings with deactivating substituents.

The folding behaviors of halide functionalized mPPEs (R = 17, 18, 19, 20) (Table 2.3) were considered separately from the other mPPEs, because of the unique nature of halide functional groups on aromatic rings. These groups have weak deactivating effects, yet still direct electrophilic substituents to *ortho* and *para* positions in aromatic substitution

reactions.[64] A distinct trend was observed in folding behaviors of halide functionalized mPPEs in solvents other than chloroform. The fluoride substituent (R = 17) showed Group 1 behavior in all five solvents, similar to mPPEs with hydrogen (R = 10) or methyl (R = 11) substituents. Yet, when changing the aromatic substituents from fluoride (R = 17) to iodine (R = 20), the behavior generally shifts towards Groups 3 and 4. This shift correlates well with the atomic weight of the halides, as do other properties of the halide functional groups, such as electro-negativities, van der Waals radii, and van der Waals interaction energies. In the OPLS force field, Lennard-Jones radius (σ) parameters for the halide substituents vary from 2.85 Å for fluoride to 3.67 Å for iodide, while the well depths (ε) increase from 0.25 kJ/mol to 2.4 kJ/mol. Thus, for the higher molecular weight halides, the position of the potential energy well becomes closer to the separation distance D_a between aromatic rings, while the energy of the interaction becomes more favorable. These effects stabilize the helical conformation and promote folding for the heavier halides.

2.5 Summary and conclusions

In this work, REMD simulations using the OPLS force field were applied as a tool for predicting the most stable conformation of numerous mPPE variants in solvents of varying polarity. For each mPPE/solvent pair, two REMD simulations were conducted, one initialized with the oligomer in an extended conformation and the other with a helical conformation, and the simulation results compared to evaluate and categorize the

respective mPPE folding behavior. The total time needed for each simulation was modest (10 ns of replica exchange dynamics), allowing us to use this procedure as a quick and simple screening test in a large scale combinatorial study. Our simulation results are in excellent agreement with available data regarding the folding behavior of mPPEs in different solvents. Further, the predictive ability of this REMD protocol was demonstrated for the folding behaviors of two newly synthesized mPPE structures in chloroform and acetonitrile (discussed in later chapter).

The REMD procedure was used to study the folding behavior of twenty uniquely functionalized mPPEs in five explicit solvents, including chloroform, acetonitrile, acetone, diethyl ether and hexane. The results were categorized into four groups, ranging from Group 1, when folding is deemed unlikely, to Group 4, when folding is deemed extremely likely. Our combined results provide evidence that polymer/solvent interactions have a significant influence over the formation of secondary structure, though folding seems to be influenced to a larger degree by the intramolecular interactions within mPPEs. Our simulations showed that electron withdrawing functional groups on the aromatic rings generally promote folding, which is consistent with previous findings. Yet, our results also predict that nonpolar solvents generally promote folding in more cases than for polar solvents, and this finding is counter to the current understanding of the folding mechanism as a process of solvophobic collapse.

Many exceptions were found to the general trends outlined above, due to the often competing effects of solvophobic/solvophilic interactions, π-stacking effects, and other

site-specific interactions. A prime example of such an exception is the folding behavior in chloroform; only one out of the twenty mPPEs in this study was found to even partially fold in chloroform, although several polymers showed clear signs of stable secondary structure in both more polar and less polar solvents. In general, mPPE secondary structure formation is not easily predicted via simple heuristics. Therefore, the complex nature of this folding phenomena and more specifically the stability of mPPE helical secondary structures is best predicted by molecular simulation (and particularly the REMD method described herein).

2.6 References

1. Heemstra, J. M.; Moore, J. S., *Org. Lett.* **2004,** *6* (5), 659-662.

2. Hill, D. J.; Mio, M. J.; Prince, R. B.; Hughes, T. S.; Moore, J. S., *Chem. Rev.* **2001,** *101* (12), 3893-4011.

3. Hill, D. J.; Moore, J. S., *Proc. Natl. Acad. Sci. U.S.A.* **2002,** *99* (8), 5053-5057.

4. Lahiri, S.; Thompson, J. L.; Moore, J. S., *J. Am. Chem. Soc.* **2000,** *122* (46), 11315-11319.

5. Nelson, J. C.; Saven, J. G.; Moore, J. S.; Wolynes, P. G., *Science* **1997,** *277* (5333), 1793-1796.

6. Prince, R. B.; Saven, J. G.; Wolynes, P. G.; Moore, J. S., *J. Am. Chem. Soc.* **1999,** *121* (13), 3114-3121.

7. Ray, C. R.; Moore, J. S., Supramolecular Organization of Foldable Phenylene Ethynylene Oligomers. In *Poly(Arylene Etynylene)S: From Synthesis to Application*, **2005**; Vol. 177, pp 91-149.

8. Stone, M. T.; Fox, J. M.; Moore, J. S., *Org. Lett.* **2004,** *6* (19), 3317-3320.

9. Stone, M. T.; Heemstra, J. M.; Moore, J. S., *Acc. Chem. Res.* **2006,** *39* (1), 11-20.

10. Yang, W. Y.; Prince, R. B.; Sabelko, J.; Moore, J. S.; Gruebele, M., *J. Am. Chem. Soc.* **2000,** *122* (13), 3248-3249.

11. Adisa, B.; Bruce, D. A., *J. Phys. Chem. B* **2005,** *109* (15), 7548-7556.

12. Adisa, B.; Bruce, D. A., *J. Phys. Chem. B* **2005,** *109* (42), 19952-19959.

13. Adisa, B. Synthesis, characterization and modeling of novel polymers and polymer-templated mesoporous materials. Ph.D. Dissertation, Clemson University, Clemson, SC, U.S.A., **2005**.

14. Elmer, S. P.; Pande, V. S., *J. Phys. Chem. B* **2001,** *105* (2), 482-485.

15. Elmer, S. P.; Pande, V. S., *J. Chem. Phys.* **2004,** *121* (24), 12760-12771.

16. Elmer, S. P.; Park, S.; Pande, V. S., *J. Chem. Phys.* **2005,** *123* (11), 114903(0-7).

17. Elmer, S. P.; Park, S.; Pande, V. S., *J. Chem. Phys.* **2005,** *123* (11), 114902(0-14).

18. Lee, O. S.; Saven, J. G., *J. Phys. Chem. B* **2004,** *108* (32), 11988-11994.

19. Tan, C. Y.; Pinto, M. R.; Kose, M. E.; Ghiviriga, I.; Schanze, K. S., *Adv. Mater.* **2004,** *16* (14), 1208-1212.

20. Nishio, M.; Hirota, M.; Umezawa, Y., *The CH-[pi] interaction : evidence, nature, and consequences.* Wiley: New York, **1998**; p 217.

21. Podeszwa, R.; Szalewicz, K., *Phys. Chem. Chem. Phys.* **2008,** *10* (19), 2735-2746.

22. Sinnokrot, M. O.; Sherrill, C. D., *J. Phys. Chem. A* **2003,** *107* (41), 8377-8379.

23. Sinnokrot, M. O.; Sherrill, C. D., *J. Am. Chem. Soc.* **2004,** *126* (24), 7690-7697.

24. Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Fujii, A., *J. Phys. Chem. A* **2006,** *110* (33), 10163-10168.

25. Arnstein, S. A.; Sherrill, C. D., *Phys. Chem. Chem. Phys.* **2008,** *10* (19), 2646-2655.

26. Lee, E. C.; Hong, B. H.; Lee, J. Y.; Kim, J. C.; Kim, D.; Kim, Y.; Tarakeshwar, P.; Kim, K. S., *J. Am. Chem. Soc.* **2005,** *127* (12), 4530-4537.

27. Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K., *J. Phys. Chem. A* **2002,** *106* (17), 4423-4428.

28. Sugita, Y.; Okamoto, Y., *Chem. Phys. Lett.* **1999,** *314* (1-2), 141-151.

29. Beck, D. A. C.; White, G. W. N.; Daggett, V., *J. Struct. Biol.* **2007,** *157* (3), 514-523.

30. Periole, X.; Mark, A. E., *J. Chem. Phys.* **2007,** *126* (1), 014903(1-11).

31. Seibert, M. M.; Patriksson, A.; Hess, B.; van der Spoel, D., *J. Mol. Biol.* **2005,** *354* (1), 173-183.

32. Zhang, W.; Wu, C.; Duan, Y., *J. Chem. Phys.* **2005,** *123* (15), 154105(0-9).

33. *High performance computing*. http:/citi.clemson.edu/hpc (accessed on 08/03/2011).

34. Bekker, H.; Dijkstra, E. J.; Berendsen, H. J. C., *Supercomputer* **1993,** *10* (2), 4-10.

35. Berendsen, H. J. C.; van der Spoel, D.; Vandrunen, R., *Comput. Phys. Commun.* **1995,** *91* (1-3), 43-56.

36. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J., *J. Comput. Chem.* **1997,** *18* (12), 1463-1472.

37. Hukushima, K.; Nemoto, K., *J. Phys. Soc. Jpn.* **1996,** *65* (6), 1604-1608.

38. Lindahl, E.; Hess, B.; van der Spoel, D., *J. Mol. Model.* **2001,** *7* (8), 306-317.

39. van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C., *J. Comput. Chem.* **2005,** *26* (16), 1701-1718.

40. van der Spoel, D.; Lindahl, E.; Hess, B.; van Buuren, A. R.; Apol, E.; Meulenhoff, P. J.; Tieleman, D. P.; Sijbers, A. L. T. M.; Feenstra, K. A.; van Drunen, R.; Berendsen, H. J. C. GROMACS USER MANUAL version 4.5.4. www.gromacs.org/documentation (accessed 08/03/2011).

41. *Materials Studio*, 4.4; Accelrys Software Inc: San Diago, CA, 2008.

42. Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R., *J. Chem. Phys.* **1984,** *81* (8), 3684-3690.

43. Parrinello, M.; Rahman, A., *J. Appl. Phys.* **1981,** *52* (12), 7182-7190.

44. Patriksson, A.; van der Spoel, D., *Phys. Chem. Chem. Phys.* **2008,** *10* (15), 2073-2077.

45. Humphrey, W.; Dalke, A.; Schulten, K., *J. Mol. Graphics* **1996,** *14* (1), 33-38.

46. Eisenberg, D.; McLachlan, A. D., *Nature* **1986,** *319* (6050), 199-203.

47. Eisenhaber, F.; Lijnzaad, P.; Argos, P.; Sander, C.; Scharf, M., *J. Comput. Chem.* **1995,** *16* (3), 273-284.

48. Jorgensen, W. L.; Madura, J. D.; Swenson, C. J., *J. Am. Chem. Soc.* **1984,** *106* (22), 6638-6646.

49. Jorgensen, W. L., *J. Phys. Chem.* **1986,** *90* (7), 1276-1284.

50. Jorgensen, W. L.; Maxwell, D. S.; TiradoRives, J., *J. Am. Chem. Soc.* **1996,** *118* (45), 11225-11236.

51. Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L., *J. Phys. Chem. B* **2001,** *105* (28), 6474-6487.

52. Price, M. L. P.; Ostrovsky, D.; Jorgensen, W. L., *J. Comput. Chem.* **2001,** *22* (13), 1340-1352.

53. Rizzo, R. C.; Jorgensen, W. L., *J. Am. Chem. Soc.* **1999,** *121* (20), 4827-4836.

54. Watkins, E. K.; Jorgensen, W. L., *J. Phys. Chem. A* **2001,** *105* (16), 4118-4125.

55. Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D., *J. Am. Chem. Soc.* **2002,** *124* (36), 10887-10893.

56. Reichardt, C., *Solvents and solvent effects in organic chemistry*. 3rd ed.; Wiley-VCH: Weinheim ; Cambridge, **2003**; p 640

57. Lide, D. R.; Kehiaian, H. V., *CRC handbook of thermophysical and thermochemical data*. CRC Press: Boca Raton, **1994**; p 518.

58. Torres, R. B.; Francesconi, A. Z.; Volpe, P. L. O., *Fluid Phase Equilibria* **2002,** *200* (1), 1-10.

59. Wohlfarth, C., Permittivity dielectric constant of liquids. In *CRC handbook of thermophysical and thermochemical data*, 91 ed.; Lide, D. R., Ed. CRC Press: Boca Raton, **2011**; pp (6)186-196.

60. Stone, M. T.; Moore, J. S., *Org. Lett.* **2004,** *6* (4), 469-472.

61. Adams, H.; Hunter, C. A.; Lawson, K. R.; Perkins, J.; Spey, S. E.; Urch, C. J.; Sanderson, J. M., *Chem.-Euro. J.* **2001,** *7* (22), 4863-4877.

62. Cockroft, S. L.; Hunter, C. A.; Lawson, K. R.; Perkins, J.; Urch, C. J., *J. Am. Chem. Soc.* **2005,** *127* (24), 8594-8595.

63. Sherrill, D. C., Computations of Noncovalent pi-interactions. In *Rev. Comput. Chem.*, Lipkowitz, K. B.; Boyd, D. B., Eds. Wiley: New York, N.Y., **2009**; Vol. 26, pp 1-30.

64. Carey, F. A.; Sundberg, R. J., *Advanced organic chemistry, Part A: Structure and Mechanisms*. 5th ed.; Springer: New York, **2007**; pp 771-824.

65. Brunsveld, L.; Prince, R. B.; Meijer, E. W.; Moore, J. S., *Org. Lett.* **2000,** *2* (11), 1525-1528.

66. Arnt, L.; Tew, G. N., *J. Am. Chem. Soc.* **2002,** *124* (26), 7664-7665.

67. Li, C. J.; Slaven, W. T.; Chen, Y. P.; John, V. T.; Rachakonda, S. H., *Chem. Commun.* **1998,** (13), 1351-1352.

68. Blatchly, R. A.; Tew, G. N., *J. Org. Chem.* **2003,** *68* (23), 8780-5.

CHAPTER 3

SYNTHESIS OF FUNCTIONALIZED *meta*-POLY(PHENYLENE ETHYNYLENE) OLIGOMERS WITH STABLE HELICAL SECONDARY STRUCTURE

3.1 Introduction

meta-poly(phenylene ethynylene)s (mPPEs) are a class of macromolecules that are able to fold into stable helical conformations under suitable conditions.[1-3] This biomimetic property could lead to many different applications for mPPEs, such as self-assembled nanostructures for sensor, drug delivery, and other biological applications.

The ability of an mPPE to form stable helical structures is influenced by a number of tunable factors. Among them, the effects of altering functional groups on the mPPE's aromatic rings, resulting in a so-called functionalized mPPE, are often the first factors to be considered in mPPE structural design. Those functional groups can stabilize the helical conformation through interactions with the solvent,[2] by strengthening π-stacking effects,[4, 5] or by introducing specific intramolecular hydrogen bonds.[6, 7] Other important factors affecting helix formation in solution, which have often been listed in the literature include the chain length of the mPPE[2, 8, 9] and the type of solvent(s) used.[10] Several studies have been conducted to elucidate the relationship between the factors mentioned and the folding behaviors of functionalized mPPEs.[2, 4, 8-10] However, because of the complex influences of those factors, accurately predicting a newly conceived mPPE's

folding ability based on the limited available experimental data is quite difficult. In the previous study, we presented a modeling procedure using replica exchange molecular dynamics (REMD) simulations as an alternative for that assessment.[5] This procedure provides a reliable and systematic method to evaluate the folding propensity of an mPPE given its functional groups and solvent conditions. These modeling results offer invaluable information that helps guide the synthesis work.

In this study, we apply this REMD procedure to examine the folding behaviors of two new functionalized mPPE structures (shown in Figure 3.1). These two mPPEs represent a new practical framework to synthesize functionalized and foldable mPPEs. In both molecules the ester and nitrile (−CN) functional groups are arranged in an alternating pattern in the *exohelix* position *meta* to the ethynylene linkages, while in mPPE1 an ether ($-OCH_3$) functional group was additionally positioned *para* to each nitrile group. Thus, when folded into a helix, both molecules have ester and nitrile groups exposed along their outer surface, while mPPE2 has the additional *endohelix* ether groups inwardly oriented within the helical cavity. These mPPEs were designed upon the notion, presented in an earlier study,[5] that the ester functional groups provide a very strong stabilizing effect on the mPPE helical structures. Thus, having those ester functional groups would favor the formation of helical mPPE secondary structure as well as enhance the solubility of the polymer, assuming that the ester groups contained ether functional groups. The other positions on the aromatic rings, such as the ether or the nitrile positions on these two mPPEs, could be replaced with a variety of different *exohelix* or *endohelix* functional groups. We further illustrate the flexibility of this mPPE concept by studying the

synthesis and folding characteristics of two new water soluble mPPEs (mPPE3 and mPPE4) and their precursors.

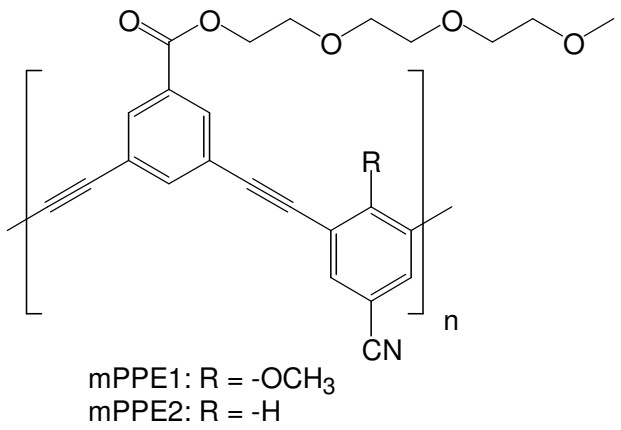


Figure 3.1 Structures of mPPE1 and mPPE2 having an alternating arrangement of *exohelix* functional groups. The *exohelix* nitrile and *endohelix* R group represent the customizable functional groups on this new mPPE backbone.

3.2 Simulation on the folding behaviors of mPPE1 and mPPE2

3.2.1 Simulation details

All simulations were conducted using the *Palmetto* supercomputer[11] at Clemson University. All of the computers used in this study were Dell PE 1950 with 2x Intel Xeon E5345 Quad Core processors at 2.33 GHz, 4MB L2 Cache, 12GB RAM and 80GB of local storage (120TB of network storage). The computers are linked to a common network by 10G-SW32LC-16M Myrinet line cards.

The REMD protocol employed in this chapter is identical to that described in earlier work[5, 12-14] and, as such, is only briefly described here. All constant volume REMD and MD simulations were conducted using Gromacs, version 3.3.1.[15-21] The starting mPPE structures were built in Materials Studio version 4.4,[22] then transferred to Gromacs. Isothermal conditions for all simulations were maintained using Berendsen temperature coupling[23] (T = 300 K, τ_T = 0.1 ps), whereas Parrinello-Rahman pressure coupling[24] (P = 1 bar, τ_P = 1 ps) was used to maintain system pressure during 200 ns isothermal-isobaric simulations that were used to equilibrate the solvated mPPE systems before REMD and MD simulations. The LINCS algorithm[17] was employed to hold constant all mPPE bond lengths. To evaluate the folding behaviors of a given mPPE system, we used Gromacs to conduct two REMD simulations, one initialized with the mPPE in a random extended structure and the other with the polymer in a helical conformation. Results were used to categorize the various mPPEs into folding property groups, based on the polymer conformations favored during simulations. As described in the previous chapter, several structural and energy parameters were used to monitor changes in polymer structure. However, only the time evolutions of the radius of gyrations (R_g) are presented in this study. The folding propensity of the polymers are designated by a number from 1 to 4, where higher numbers indicate a greater likelihood that the mPPE will form a stable helix in the given solvent (i.e., Group 1 being least likely and Group 4 being most likely).

For this study, we employed the simplified models of mPPE1 and mPPE2, similar to those presented in earlier studies.[5, 12-14] The long ether pendants, because they do not play an important role in helix formation in this simulation study, were replaced by methyl

groups as done earlier. The mPPE and solvent models were taken from the Optimized Potentials for Liquid Simulations (OPLS),[25-31] following the procedure described earlier.[5, 12-14] Two explicit solvents, acetonitrile and chloroform with the same parameters as described earlier,[5] were used for this study.

3.2.2 Modeling prediction of mPPE1 and mPPE2 folding behaviors

The folding behaviors of mPPE1 and mPPE2, each having six repeat units (12 aromatic rings), were simulated and measured in chloroform and acetonitrile. REMD simulation results showed that mPPE1 and mPPE2 were both categorized in Group 1 in chloroform and Group 3 in acetonitrile (Figure 3.2), using the categorization presented in the previous study.[5] Thus, these mPPEs should not fold in chloroform, but should at least partially fold in acetonitrile.

Compared to the folding capability of the ester mPPE in acetonitrile (Group 4), these results suggest that reducing the number of ester functional groups does decrease the propensity of the mPPEs to fold from their initial extended conformations into helical conformation. The data also suggest that the helical stability is not enhanced by the presence of the nitrile or ether functional groups on the mPPE back bone, in agreement with the modeling results and experimental results,[4, 5] which showed that for mPPEs having only nitrile or ether functional groups exhibited little tendency to form helical structures in acetonitrile (Group 2).

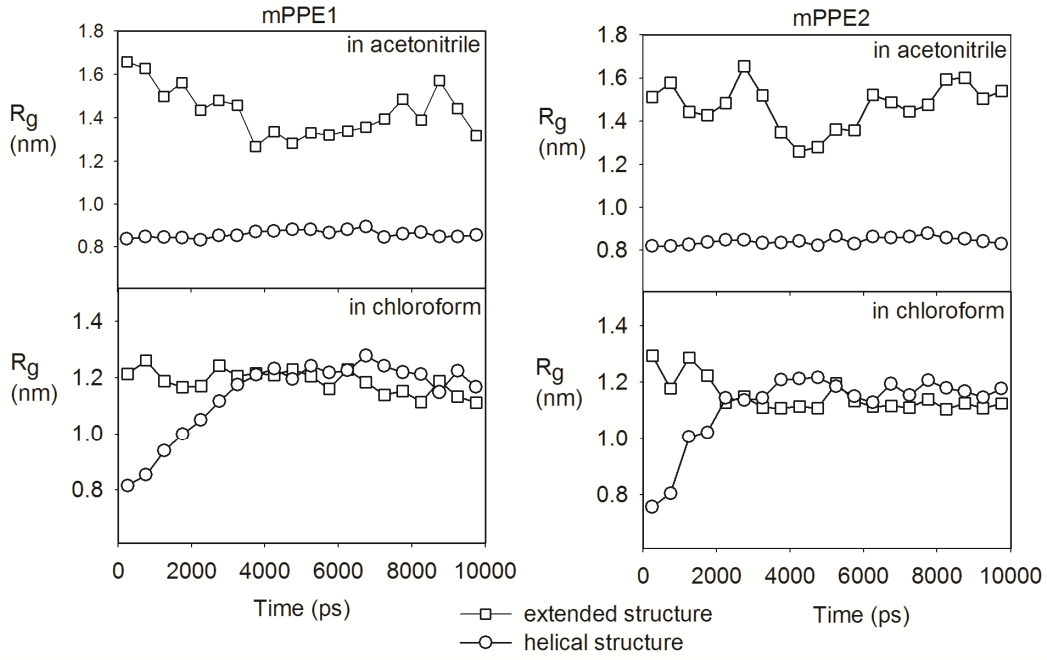


Figure 3.2 Time evolution of the radius of gyration (R_g) of mPPE1 and mPPE2 in acetonitrile and chloroform. Squares are from simulations beginning with an extended structure, and circles are from simulations beginning with a helical structure. Each plotted data point is a time average over 500 ps, placed at the center of the corresponding time interval.

To compensate for the reduction in stabilization due to the smaller number of ester functional groups, one possible approach is to increase the chain length of these mPPEs. This follows from observation that, generally, longer mPPEs have more stable helical structures in solution.[2, 9] Because of the 6-mer chain length of mPPE1 and mPPE2, they could at least partially coil up in acetonitrile (Group 3 behavior); however, a higher number of repeat units would likely allow them to fully fold, falling under our designation of Group 4, since there are more π-stacking aromatic pairs to initialize the

folding process and stabilize the helical intermediates. To explore this approach, additional REMD simulations were performed in explicit acetonitrile for the extended structures of mPPE1 and mPPE2 having chain lengths of 7 and 8 repeat units (14 and 16 aromatic rings, respectively). Only the simulations beginning with an extended structure were conducted, since the previous simulations have indicated that the 6-mer helical structures would be stable in acetonitrile. The REMD simulation results clearly demonstrated that both extended structures of the 7-mer and 8-mer mPPEs would be able to fold into stable helical structures in acetonitrile (Figure 3.3).

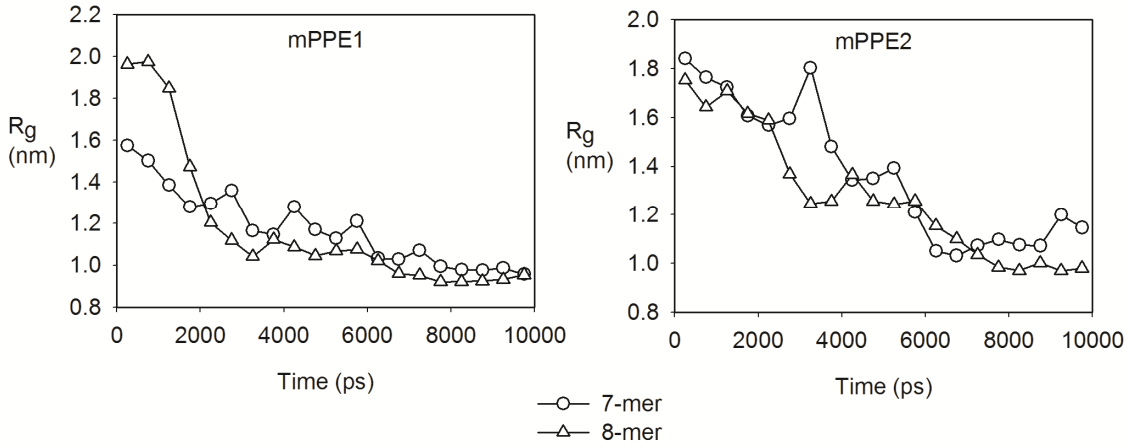


Figure 3.3**.** Time evolution of the radius (R_g) of gyration for long chain length mPPE1 and mPPE2 initialized as extended structures in acetonitrile. Circles are from simulations of 7-mer mPPEs, triangles are from simulations of 8-mer mPPEs. Each plotted data point is a time average over 500 ps, placed at the center of the corresponding time interval.

3.3 Synthesis and folding characterization of mPPE1 and mPPE2

3.3.1 Synthesis route and folding characterization using UV absorbance and fluorescence emission spectroscopies

Because the models predict that mPPE1 and mPPE2 could potentially fold into stable helical conformations in acetonitrile, the actual polymers were produced using the synthesis route, similar to those published earlier (Figure 3.4).[32-34] The folding behavior of each mPPE sample was evaluated using UV absorbance and fluorescence emission spectroscopy.

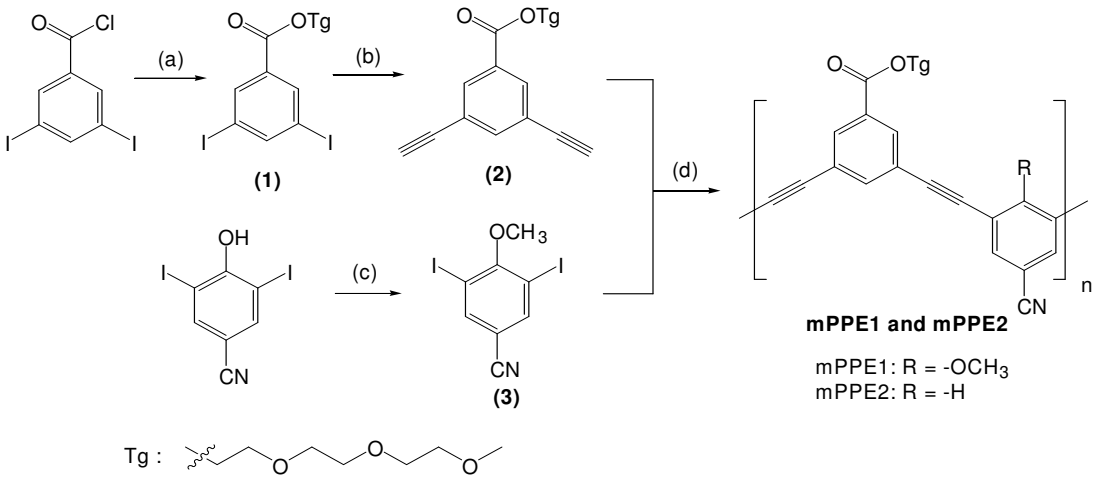


Figure 3.4 Synthesis route for mPPE1 and mPPE2. Reagents: (a) 2-(2-(2-methoxyethoxy)ethoxy)ethanol (TgOH), CH_2Cl_2, triethylamine (TEA), r.t, 24 h; (b) trimethylsilylacetylene (TMSA), $Pd_2(dba)_3$, CuI, $P(Ph)_3$, diisopropyl amine (DIPA), toluene, 78 °C, 24 h; then tetra-*n*-butylammonium fluoride (TBAF), THF, 2h; (c) CH_3OH, $P(Ph)_3$, diisopropyl azodicarboxylate (DIAD), r.t., 24 h and (d) $Pd_2(dba)_3$, CuI, $P(Ph)_3$, DIPA, toluene, 78 °C, 24 h.

The UV absorbance and fluorescence emission spectra for the synthesized mPPEs show that these polymers are in a fully folded state (helical) in acetonitrile and disordered (not folded) in chloroform (Figure 3.5). As shown in previous studies,[2] the UV absorbance at 305 nm (mPPE1) and 307 nm (mPPE2) decreased upon changing the solvent from chloroform to acetonitrile, thus, indicating more *cisoid* (helical) conformations. Further, with the fluorescence spectra, the absence of a peak at 380 nm (mPPE1) and 350 nm (mPPE2) and the presence of a peak at 425 nm (mPPE1) and 450 nm (mPPE2) can be taken as indicators of a helical structure in acetonitrile.[2]

The adsorption and emission data clearly indicate that both mPPEs in acetonitrile are experimentally categorized as Group 4 systems or fully folded (helical). The results also imply that they may have longer chain lengths, at least 7-mer, as suggested from REMD simulations.

The combination of modeling and experimental results indicates that the helical conformations of the two new mPPEs are possibly less stable than that of the ester mPPE at comparable chain lengths, and that the synthesized mPPEs likely have more than six repeat units (12 aromatic rings). To quantify those assertions, gel permeation chromatography (GPC) and denaturation titration analyses[8] were conducted. From GPC analysis, the number of repeat units in each mPPE sample could be estimated. From the denaturation experiments, the free energy difference between the folded and unfolded states in acetonitrile ($\Delta G_{folding,CH_3CN}$) and the volume percentage of chloroform in the

mixture ($C_{CHCl_3,50}$) at which 50% of mPPE molecules are in an unfolded state were calculated.

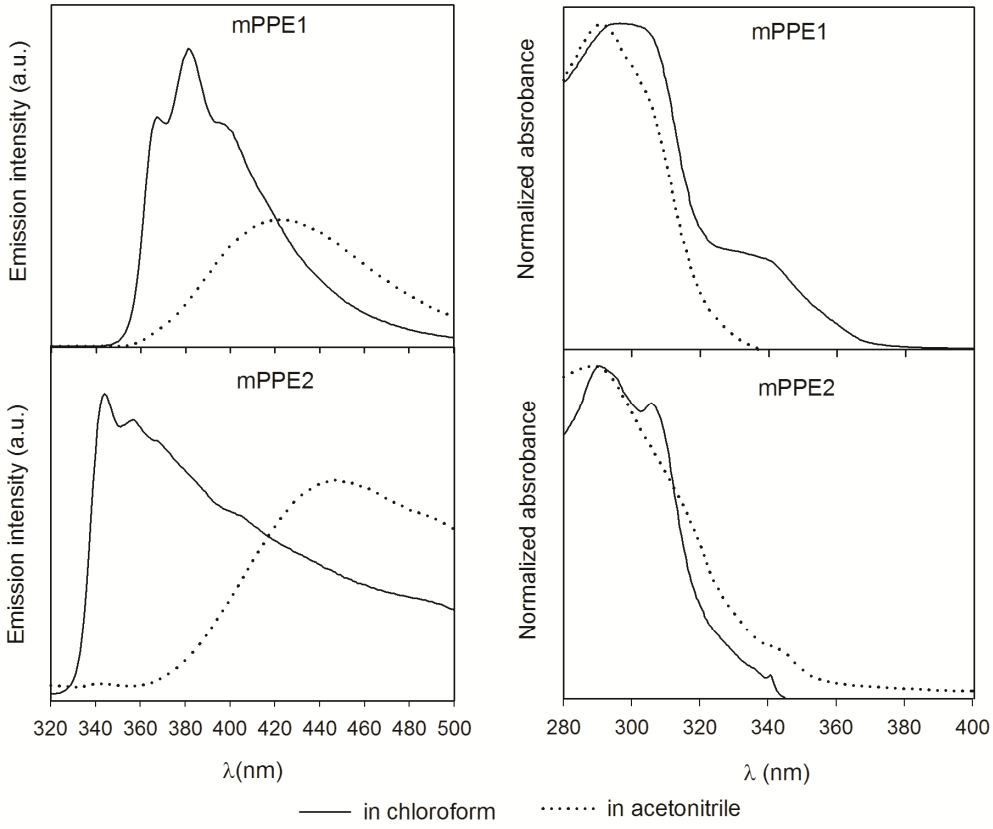


Figure 3.5 Fluorescence emission and UV absorbance (normalized) spectra as measured for mPPE1 and mPPE2 in chloroform (solid lines) and acetonitrile solvents (dashed lines).

3.3.2 Chain length estimation using gel permeation chromatography experiments

The accurate estimation of polymer chain length using gel permeation chromatography (GPC) would require the use of a series of functionalized mPPEs with precise chain lengths as standards. The other GPC methodologies available to precisely estimate polymer chain length involve using a series of polystyrene standards and an additional viscometer detector; or a single standard and triple detectors including light scattering, viscometry and refractive index.[35-39] Our GPC data, however, using polystyrene having narrow molecular weight distribution as the standard on a GPC instrument equipped with only refractive index and UV absorbance detectors, and with chloroform as solvent to ensure that mPPE was in uncoil conformations, requiring an indirect method for analyzing the raw experimental results. Huang and Tour[40] and Bunz[41] showed that for the *para*-poly(phenylene ethynylene) (pPPE), the actual molecular weight of the polymer sample is overestimated roughly by a factor of two in GPC results using polystyrene narrow molecular weight distribution standards. From that, the chain length of the synthesized mPPEs could be estimated using the following equation:

$$\text{Estimated number of aromatic rings} = 2\times\left(\frac{M_p}{F\times M_{repeatunit}}\right) \qquad (3.1)$$

where M_p is the peak molecular weight of the studied polymer from GPC, $M_{repeatunit}$ is the calculated molecular weight of its repeat unit and F is the overestimation factors from using polystyrene narrow molecular weight distribution as the standards.

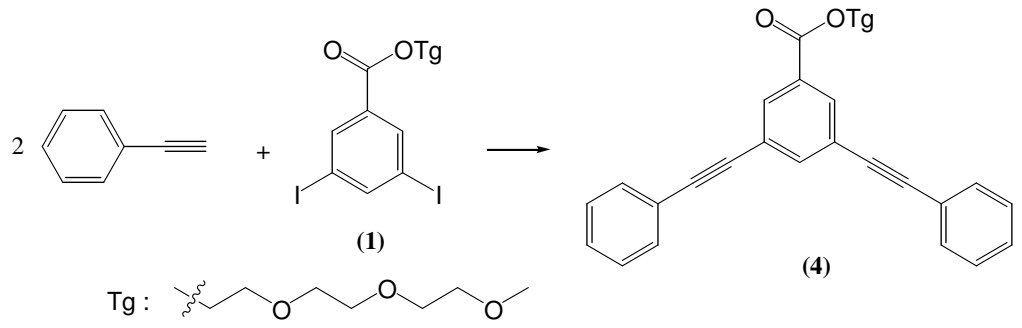


Figure 3.6 Reaction used to synthesize an mPPE trimer analog molecule, mPPEa. Reagents: $Pd_2(dba)_3$, CuI, $P(Ph)_3$, diisopropyl amine (DIPA), toluene, 78 °C, 24 h.

Because pPPEs have linear back bones, and mPPE aromatic rings are connected at *meta* positions, the pPPEs are less able to form compact secondary structures as compared to mPPEs, which means estimation using Eqn. (3.1) with F = 2 would underestimate the actual mPPE chain lengths. This is based on the assumption that an mPPE oligomer will have smaller hydrodynamic radius as compared to a like sized pPPE because of its flexible polymer back bone.[42] Therefore, mPPE chain lengths calculated via Eqn. 3.1 are only meant to serve as a lower bound for the actual mPPE chain length. This simple method was further refined by adjusting the value of the overestimation factor F in the equation by synthesizing of a short chain mPPE analog, mPPEa, with 3 aromatic rings (Figure 3.6), and measuring its molecular weight using GPC at the exact same conditions used for other mPPE analyses (with the same set of polystyrene narrow molecular weight distribution standards). The ratio between the molecular weight derived from the GPC data and the actual molecular weight of the mPPEa was used to calculate as the over estimation factor (F) for mPPE oligomers. The estimates in Table 3.1 show that the

number of aromatic rings in mPPE1 and mPPE2 are at minimum 20 (and likely higher). This observation supports the modeling prediction that mPPEs having more than 12 aromatic rings are capable of forming stable helical structures in acetonitrile.

Table 3.1 Gel Permeation Chromatography based estimation of polymer chain length for mPPE1 and mPPE2.

Calculated values	mPPE1	mPPE2
M_w [a]	15705	33926
M_n [a]	8113	9261
M_p [a]	8970	17929
Estimated number of aromatic rings [b]	20	42
Estimated number of aromatic rings [c]	28	60

[a] Calculated from GPC data using polystyrene narrow molecular weight distribution standards.
[b] Estimated using Eqn. (3.1) with F = 2, round off to the nearest even number.
[c] Estimated using Eqn. (3.1) with F calculated from GPC data of mPPEa, round off to the nearest even number.

$F = M_p/M_{mPPEa} = 667/468.5 = 1.4$

M_p = Molecular weight of mPPEa derived from GPC data, Da.

M_{mPPEa} = The exact molecular weight of mPPEa, Da.

3.3.3 Denaturation titration experiments for mPPE1 and mPPE2

The fluorescence emission spectra of mPPEs in several solvent mixtures of acetonitrile and chloroform were measured to determine the difference in free energy between the folded (helical) and unfolded (disordered) states (ΔG) for polymers mPPE1 and mPPE2

using the method proposed by Prince et al.[8] The following section provides a brief description of that method.

3.3.3.1 Denaturation titration method[8]

The mPPE folding reaction can be described by a simple two-state model where the polymer exists in either the folded or unfolded state. Further, the relative concentration of the two polymer states can be described by an equilibrium constant, K_{eq}, for the conformational restructuring process:

$$\text{disordered polymer conformations} \overset{K_{eq}}{\rightleftharpoons} \text{helical polymer conformation}$$

The change in free energy for this process is then given by:

$$\Delta G_{folding} = -RT\ln K_{eq} \qquad (3.2)$$

where $\Delta G_{folding}$ is the free energy difference for between folding states of an mPPE in a given solution; K_{eq} is the equilibrium constant of the mPPE folding reaction.

Prince et al.[8] has shown that, in the denaturation titration experiment, the solvents should be selected so that one initiates folding, while the other inhibits helix formation. Thus, for these mPPE systems, chloroform was chosen as the helix inhibitor solvent as it leads to 100 percent of mPPE molecules being in extended conformations, whereas acetonitrile was selected as the helix biasing solvent because it causes 100 percent of mPPE molecules to fold into helical conformation. The free energy change associated with the

helix formation in each solvent mixture has been shown to be linearly dependent with the percentage of chloroform, using the following equation:

$$\Delta G_{folding,solvent} = \Delta G_{folding,CH_3CN} - m \times C_{CHCl_3} \quad (3.3)$$

where $\Delta G_{folding,solvent}$ is the free energy difference for between folding states of an mPPE in a given solution; $\Delta G_{folding,CH_3CN}$ is the free energy difference between mPPE folding states in acetonitrile; m is a linearized coefficient; C_{CHCl_3} is the concentration of chloroform in solution as a volume percentage (v/v).

From Eqn. (3.3), the free energy difference of folding in acetonitrile solvent, $\Delta G_{folding,CH_3CN}$, is calculated by finding the concentration of chloroform at which the $\Delta G_{folding,solvent}$ of folding in solvent equals nil, or there is an equal amount of folded and unfolded states:

$$\Delta G_{folding,CH_3CN} = m \times C_{CHCl_3,50} \quad (3.4)$$

where $C_{CHCl_3,50}$ is the volume percentage of chloroform in an acetonitrile mixture that leads to 50% of the mPPE molecules being in an unfolded state.

The $\Delta G_{folding,solvent}$ can be calculated from the equilibrium constant for mPPE helix formation in a given solvent mixture using Eqn. (3.2), and the appropriate equilibrium constant can be found knowing the concentration of the mPPE polymer that exists in either helical or disordered states:

$$K_{eq} = \frac{F_{helix}}{F_{disordered}} = \frac{1 - F_{disordered}}{F_{disordered}} \quad (3.5)$$

where F_{helix} is the fraction of polymers in a helical conformation.

The F_{helix} fraction can be found from experimentally measured fluorescence data using the equation:

$$F_{disordered} = \frac{I_A - I_{solvent}}{I_A - I_C} \quad (3.6)$$

where $F_{disordered}$ is the fraction of mPPEs in an unfolded conformation in solution; I_C is the emission peak intensity for the mPPE in pure chloroform; I_A is the emission spectra intensity at the same wavelength as I_C but for the mPPE in pure acetonitrile; $I_{solvent}$ is the emission spectra intensity at the same wavelength as I_C but for the mPPE in a given mixture of acetonitrile and chloroform.

3.3.3.2 Denaturation titration experiment for mPPE1 and mPPE2

To evaluate the stability of the helical structures of mPPE1 and mPPE2, as well as to estimate the number of aromatic rings in each polymer sample, the denaturation titration experiments using mixtures of chloroform and acetonitrile as solvent were conducted for mPPE1 and mPPE2 samples as described in the previous section. We note that the broad molecular weight distributions of mPPE1 and mPPE2 samples could have some effect on the titration calculation, because there is short-chain length fraction of the polymer

sample which does not fold in acetonitrile. The results are given in Figure 3.7 and summarized in Table 3.2.

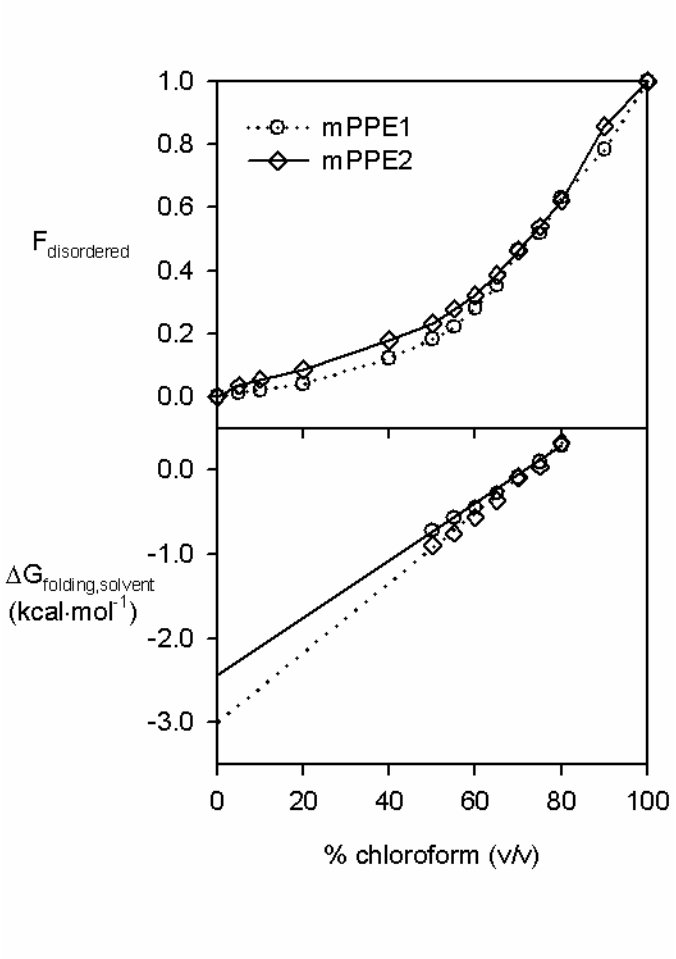


Figure 3.7 Plots of the fractions of mPPE1 and mPPE2 that exist in an unfolded conformation, $F_{disordered}$ as calculated using Eqn. (3.5) and $\Delta G_{folding,solvent}$ for mPPE folding calculated from Eqn. (3.2) and Eqn. (3.5) as a function of chloroform volume percentage. The data are from denaturation titration experiments using chloroform/acetonitrile mixtures as solvent.

Table 3.2 Denaturation titration data for mPPE1 and mPPE2.

Calculated values	mPPE1	mPPE2
$C_{CHCl_3,50}$ [a]	73	72
$\Delta G_{folding,CH_3CN}$ (kcal·mol^{-1})[b]	-3.0	-2.4

[a] Volume percentage of chloroform in an acetonitrile mixture at which 50% of the mPPE molecules are in an unfolded state. Interpolated using $F_{disoredered}$ data in Figure 3.6.

[b] Extrapolated using $\Delta G_{folding,solvent}$ data in Figure 3.6 at $C_{CHCl_3} = 0$.

The plots above show that for both the mPPE1 and mPPE2 samples, the percentage of oligomers that are disordered, $F_{disordered,}$ gradually increases with increases in the concentration of chloroform in solution (the linearized coefficient *m* in Eqn. (3.3) is 40.9 and 33.8 cal/mol for mPPE1 and mPPE2, respectively). This behavior differs from that observed with titration experiments for the ester mPPE reported by Prince et al.,[8] which showed more dramatic changes in folding behavior with the addition of chloroform (e.g., m = 50 and 99 cal/mol for 12-mer and 18-mer ester mPPEs, respectively). These differences may be attributed to mPPE1 and mPPE2 having a broader molecular weight distribution than the reported ester mPPEs, which have precisely controlled chain length. Compared to the ester mPPE, whose 18-mer structure (having 18 aromatic rings) has a $\Delta G_{folding,CH_3CN} = -7.1$ kcal·mol^{-1}, the calculated $\Delta G_{folding,CH_3CN}$ and the estimated number of aromatic rings for mPPE1 and mPPE2 reveal that their helical structures in acetonitrile

are less stable than that of the ester mPPE having the same chain length, agreeing with assertions derived from REMD simulations, which showed that the smaller numbers of ester functional groups may lessen the mPPE folding propensity.

3.4 Synthesis of water soluble amine functionalized mPPEs

A considerable limitation of many functionalized and unfunctionalized mPPEs is that they are insoluble in water, which precludes their use in biological applications. Further, for the few mPPEs that are reported to be soluble in water or water-rich solutions, not all of them are observed to fold into a helical conformation once solvated. For example, Arnt and Tew[32, 33] reported an mPPE that is soluble in a mixture of DMSO and water, but did not observe any evidence of a stable helical conformation in that solution. Li et al.[43] reported an mPPE with acid functional groups that exhibits a gel like property in water solutions. However, examples of mPPEs that are both water soluble and folding biased are quite rare. Stone et al.[44] reported an mPPE with long ether pendants, which stabilize the helical conformation and make the polymer soluble in water, and finally, Tan et al.[45] reported an ionic mPPE containing sulfate groups, which folded into a helical structure in water.

In an effort to expand the number of known water soluble mPPEs, we report the synthesis of two new mPPE structures that are both highly soluble in water and able to form stable helical structures in aqueous solutions. It is also significant that these polymers contain

readily accessible amine functional groups, which may prove useful for biological applications that require specific and directed interactions with enzymes or proteins. Additionally, we report the folding behaviors of these two mPPEs and their precursors in several other protic and aprotic solvent systems.

3.4.1 Synthesis route and chain length evaluation

The polymers were synthesized using the synthesis route depicted in Figure 3.8. This procedure is similar to those reported earlier.[32-34] The GPC results for the protected mPPEs are in Table 3.3, and the numbers of aromatic rings in the mPPE samples were estimated as in Section 3.3. Two samples of the mPPE4 variation (R = H) were obtained via flash column chromatography, and we denote these polymers as mPPE4L (long chain length polymer) and mPPE4S (short chain length polymer). This allowed us to examine the chain length dependence of secondary structure formation in the polymer, which has previously been shown to be significant.[8, 9] We note that small amounts (less than 5%) of very high molecular weight polymer were present in the p_mPPE4 and p_mPPE4L samples, but were not included in molecular weight averages because they were beyond the range of the GPC calibration standards. The lower bounds listed in Table 3.3 showed that the mPPE samples were estimated to be longer than the seven repeat units needed for folding.[2, 9]

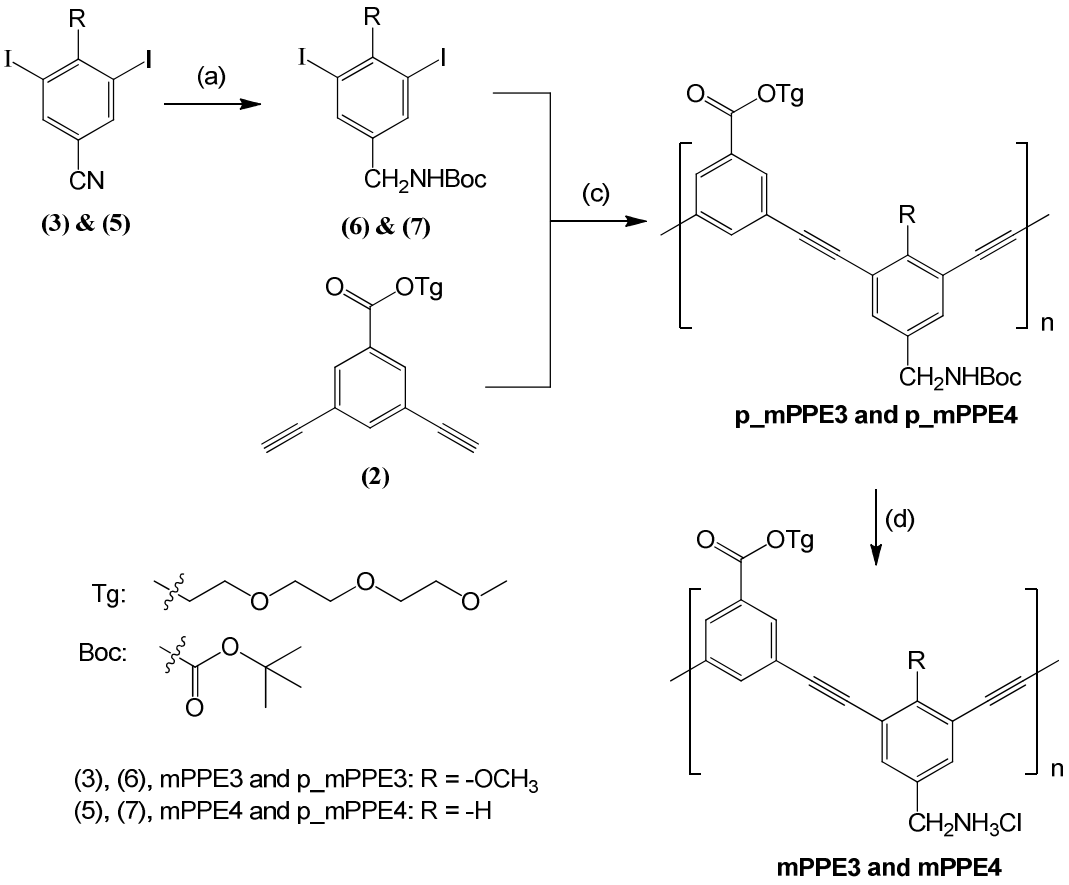


Figure 3.8 Synthesis route for mPPE3 and mPPE4. Reagents: (a) BH_3·THF, THF, reflux, 24 h; then $(Boc)_2CO$, N,N-dimethylformamide (DMF), NaOH, H_2O, 24 h; (b) $Pd_2(dba)_3$, CuI, $P(Ph)_3$, diisopropyl amine (DIPA), toluene, 78 °C, 24 h; (c) HCl 4M/dioxane, CH_2Cl_2, 0 °C, 2 h.

Table 3.3 Estimated numbers of aromatic rings using GPC data for p_mPPE3, p_mPPE4L and p_mPPE4S.

Calculated values	p_mPPE3	p_mPPE4L	p_mPPE4S
M_w[a]	21735	12591	8376
M_n[a]	9622	6822	3818
M_p[a]	9481	8661	3812
Estimated number of aromatic rings[b]	18	16	8
Estimated number of aromatic rings[c]	26	24	10

[a] Calculated from GPC data using polystyrene narrow molecular weight distribution standards.
[b] Estimated using Eqn. (3.1) with F = 2, round off to the nearest even number.
[c] Estimated using Eqn. (3.1) with F = 1.4, calculated from GPC data of mPPEa as in Section 3.3.2, round off to the nearest even number.

3.4.2 Folding characteristic

The folding behavior of each mPPE sample was evaluated using ultraviolet (UV) and fluorescence spectroscopy.[2, 4, 5, 10] Specifically, the folding behaviors of protected amine polymer (p_mPPE3, p_mPPE4L and p_mPPE4S) were determined in acetonitrile and chloroform, and the folding behaviors of the amine functionalized mPPEs (mPPE3, mPPE4L, and mPPE4S) were determined in ethanol, methanol, 1-propanol, 2-propanol, and water. For mPPE3 and the protected amine precursor p_mPPE3, the UV absorbance ratio A_{305}/A_{289} and the fluorescence emission ratio I_{450}/I_{380} were examined (subscripts denote wavelengths in nm). However, for mPPE4L, mPPE4S, and their protected amine

precursors, p_mPPE4L, and p_mPPE4S, slightly different absorbance (A_{305}/A_{290}) and emission (I_{425}/I_{350}) ratios were used to evaluate secondary structure formation.

Several features of the UV spectra made this method less effective in determining the folding behaviors of our novel mPPE samples as compared to those reported in previous studies.[2, 4] For example, the absorption spectra for mPPE3 and p_mPPE3, shown in the insets of Figure 3.9a and Figure 3.9d, are generally broad and lack the characteristic shoulder in the vicinity of 305 cm^{-1}, preventing a reliable assessment of the amount of *cisoid* conformations present in the mPPEs. Further, when examining the deprotected mPPEs, the shift of the maximum absorbance wavelength in water makes it difficult to compare the amount of *cisoid* conformations relative to the other solvents based on absorbance at fixed wavelengths (i.e., using the A_{305}/A_{289} ratio). Because of these features, the emission ratios from fluorescence emission spectra were used as the primary tool for characterizing the folding behavior of each mPPE system, with the absorbance ratio from UV absorbance spectra playing a supplementary role. The fluorescence emission and UV absorbance ratios are summarized in Table 3.4. A fluorescence emission ratio greater than 1 indicates that the mPPE exists as helical conformation in solution.

Table 3.4 Secondary structure formation by mPPE samples in different solvents as determined by fluorescence emission ratio and UV absorbance ratio.

Solvents	Fluorescence emission ratio[a]			UV absorbance ratio[b]		
	p_ mPPE3	p_ mPPE4L	p_ mPPE4S	p_ mPPE3	p_ mPPE4L	p_ mPPE4S
Chloroform	0.14	0.39	0.16	0.96	0.85	0.83
Acetonitrile	0.34	9.70	0.29	0.74	0.71	0.74
Methanol	0.34	11.29	0.54	0.74	0.69	0.73
Solvents	mPPE3	mPPE4L	mPPE4S	mPPE3	mPPE4L	mPPE4S
Methanol	0.23	0.73	0.17	0.86	0.87	0.82
Water	5.90	7.73	4.30	0.69	0.68	0.62
Ethanol	0.16	2.61	0.20	0.85	0.72	0.74
1-Propanol	0.15	3.21	0.17	0.77	0.70	0.77
2-Propanol	0.12	1.83	0.59	0.86	0.80	0.86

[a] Reported emission ratios for mPPE3 and p_mPPE3 are I_{450}/I_{380}; for all other samples is I_{425}/I_{350}.

[b] Reported absorbance ratios for mPPE3 and p_mPPE3 are A_{305}/A_{289}; for all other samples is A_{305}/A_{290}.

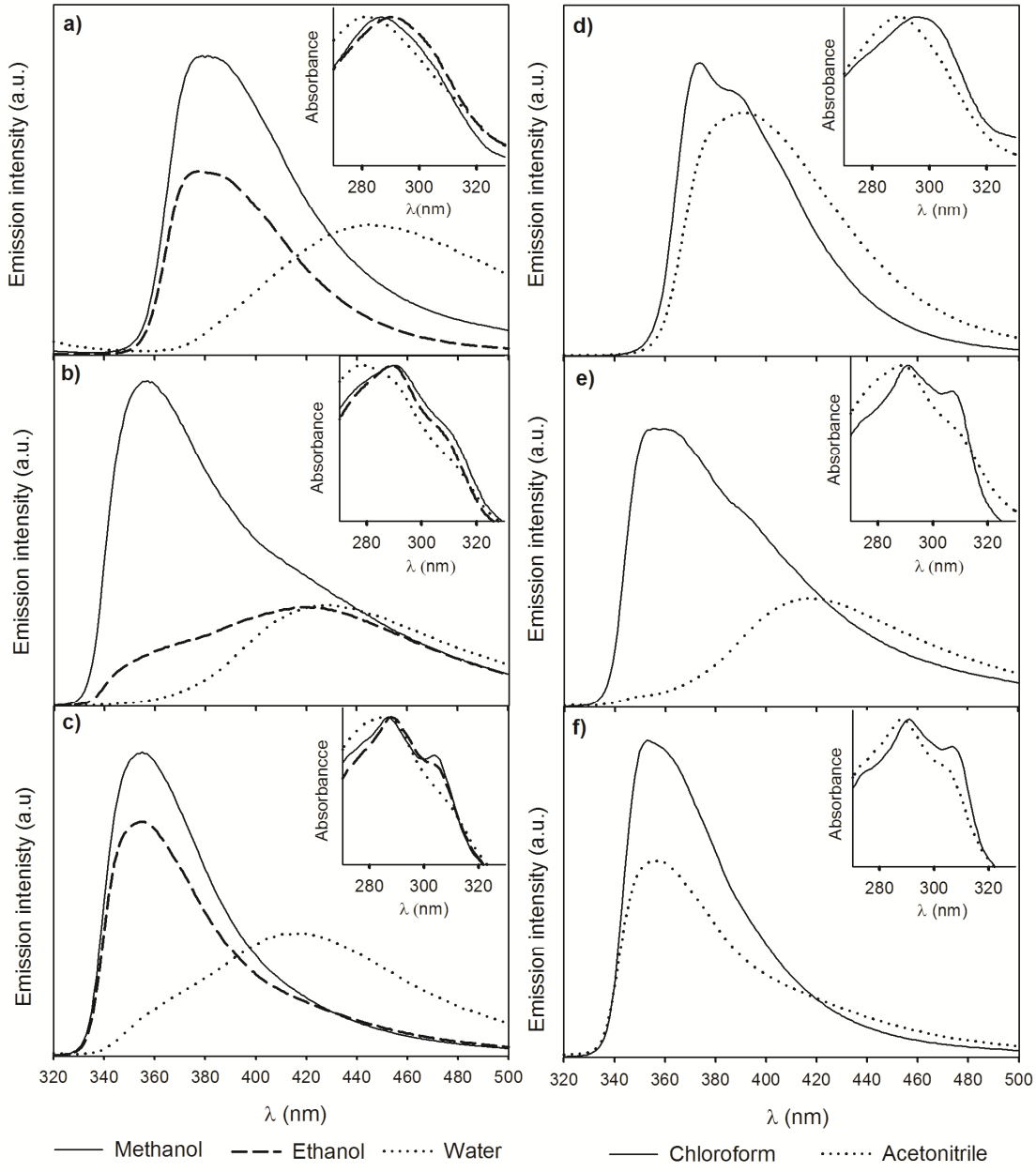


Figure 3.9 Fluorescence emission and UV absorbance (normalized, inset) spectra of the studied mPPE systems: a) mPPE3, b) mPPE4L, and c) mPPE4S in methanol, ethanol and water, and d) p_mPPE3, e) p_mPPE4L, and f) p_mPPE4S in chloroform and acetonitrile.

3.4.3 Factors affecting the folding behaviors of newly synthesized functionalized mPPEs

Folding behavior was evaluated for the mPPE samples spanning a range of three independent variables: solvent type, mPPE chain length, and mPPE chemical functionality (as determined by the functional groups R listed in Figure 3.8). Results in previous sections and earlier studies have established that each of these factors can influence the stability of the mPPE helical conformation,[2, 4, 9] and in general, because of the complex inter-relationship among those factors, no simple trend can be precisely anticipated with respect to them without the use of sophisticated molecular simulations.[5] A combinatorial approach was used in this work, the results of which are presented in Figure 3.9 and summarized in Table 3.4. The following sections present the effects of each independent variable on the folding of the mPPE samples.

3.4.3.1 Solvents effects on mPPE helical structure formation

The deprotected mPPE samples synthesized in this study were soluble in water and other protic solvents. Both of the mPPE variations (mPPE3 and mPPE4) were found to form a helical secondary structure in aqueous solution, making them viable candidates for biological applications. Additionally, the mPPE4L sample exhibited stable helical structures in ethanol, 1-propanol, and 2-propanol solutions, though the other deprotected samples did not show evidence of folding in solvents other than water. Finally, the folding behaviors of all of the mPPE samples dissolved in methanol were characterized,

but the protected polymer p_mPPE4L was the only one to exhibit an ability to fold into helical conformations.

While some solvents are known to have a strong effect on the folding behavior of mPPEs, others are rather unpredictable in this respect. For example, chloroform largely inhibits helix formation by mPPEs, while some mPPEs in acetonitrile fold into a helix and others do not.[2, 5, 10] Thus, in addition to the protic solvents considered in this study, several other solvents were used, which have previously been shown to promote or hinder folding. We found that chloroform inhibits helix formation by the protected mPPE samples (p_mPPE3 and p_mPPE4), which is consistent with the solvent's known propensity to denature the helical conformation of mPPEs. Similarly, no ordered folding was observed for any of the reported samples in acetonitrile solution, except for the long chain length protected polymer sample p_mPPE4L.

3.4.3.2 mPPE chain lengths effect on helical structure formation

The experimental results indicate a significant effect of the polymer chain length on the stability of the mPPE2 helical structures. While both mPPE4L and mPPE4S folded into an ordered helical conformation in water, the long chain length mPPE variants (mPPE4L and p_mPPE4L) were found to form stable, helical secondary conformations in several other solvents, including acetonitrile, methanol, ethanol, and propanol. Based on this data, we speculate that there exists a minimum chain length for mPPE4, below which a

given oligomer will not fold into a helical conformation even in favorable solvent conditions.

Those conclusions agree with simulation results in acetonitrile for mPPE1 and mPPE2 in Section 3.2. While the 6-mer polymers are predicted to partially fold into helical conformations (Group 3 behavior)), the 7-mer and 8-mer mPPEs are predicted to become fully folded (Group 4 behavior). Our results are also consistent with those of Stone et al. ,[9] which indicated there was a discernible change in the stability of the helical conformation with respect to chain length for ester-functionalized mPPEs. In their study, the onset of helical stability was shown to occur at or near twelve repeat units, and that stability improved with increasing chain length. The GPC based estimations of mPPE chain length for oligomers examined in this study are listed in Table 3.3 and indicate that all three mPPE samples have a sufficient number of aromatic rings to stabilize a helical conformation should that structure be favored under the specific solvent conditions. Because there is evidence of a stable helical structure for mPPE4S in aqueous solution, the p_mPPE4S and mPPE4S samples must be at least seven repeat units long (and are likely significantly longer), as this is the minimum number of repeat units needed to form a single helical turn with one π-stacking interaction (data shown in Table 3.3), in agreement with the estimation using overestimation factor calculations. Further, the GPC data shown in Table 1 clearly indicates that the p_mPPE4S and mPPE4S oligomers are of sufficient length to exhibit stable helical structures. Thus, helix formation is possible for mPPE4S, but the spectroscopy data indicate that the polymer is not stable in alcohol at

that chain length. The mPPE4L sample, on the other hand, contains sufficiently long chains to stabilize the helical structure in such solvents.

3.4.3.3 Impact of functional groups on secondary structure formation

By comparing the folding behaviors of mPPE3 and mPPE4, as well as their protected amine precursors, we can deduce the effect of the primary structure of the mPPE copolymers on the stability of their respective helical conformations. Structurally, the only difference between the two materials is that mPPE3 contains a methoxy ($-OCH_3$) functional group at each position *para* to the amine substituents, whereas mPPE4 has a hydrogen in each of these positions. Further, these functional groups (R) are positioned such that they would lie in the interior regions of the helix (i.e., they are *endohelix* functional groups) when the mPPE exhibits a helical secondary structure, which would significantly limit their interaction with the solvent. Despite this fact, the experimental results showed that this simple change from a methoxy group to a hydrogen atom can have a significant effect on the folding behavior of the respective mPPE. Although GPC results indicate the chain length of the methoxy functionalized polymer (mPPE3) is greater than that of the unfunctionalized mPPE4L, the latter oligomer was found to fold into a helical conformation in nearly all tested alcohol solvents, while mPPE3 remained amorphous in all of them. Similarly, when comparing the precursors p_mPPE3 and p_mPPE4L, the sample without methoxy substituents exhibited a propensity to fold into helical structures in acetonitrile and methanol, whereas the methoxy-containing polymer (p_mPPE3) did not.

The results described above are generally consistent with the conventional view of folding as a process of solvophobic collapse. In the helical conformation of mPPE3, the methoxy groups point towards the interior of the helical cavity and are shielded from the solvent molecules. In its unfolded state, the methoxy groups are exposed to the solvent. Thus, in polar solvents, the unfolded state of mPPE3 is stabilized by favorable solvent interactions. When the methoxy group is replaced with a hydrogen atom, as in mPPE3L, the site becomes less-polar and therefore, has less favorable interactions with polar solvent molecules. This leads to the solvophobic collapse of mPPE4L into a helix to minimize these interactions.

The above explanation is consistent with the results in Table 3.4 for the mPPEs in alcohol solvents. Yet, a puzzling contradiction is given by the data in aqueous solution. Because water is more polar than the alcohol solvents, one would expect it to stabilize the unfolded or random conformations of mPPE3. According to our data, however, mPPE3 shows evidence of folding in aqueous solutions. This could also be resulted from entropic effect of breaking water hydrogen bond network, so that those systems favor mPPEs in helical conformations. This result proves that the relationship between primary and secondary structure in mPPEs is nontrivial.

Despite the difference in the folding behaviors for the long chain length samples mPPE3 and mPPE4L, we found that the folding behavior of mPPE3 was very similar to the lower molecular weight sample mPPE4S. These results are important given the previously discussed results regarding chain length dependence. It is most likely the case that

mPPE3, like mPPE4S, does not have sufficient molecular weight (i.e., chain length) to form stable helical structures. Further, the onset of helix stability, mentioned in the previous section, is possibly higher for mPPE4 than for mPPE3.

In addition to the solvophobic collapse arguments discussed previously, there is evidence that the ether groups attached to the aromatic rings (as present in mPPE3) can further destabilize the helical conformation by disrupting π-stacking interactions. Lahiri et al.[4] showed that mPPE macrocycles were unable to agglomerate when they contained ether substituents, and the corresponding mPPE oligomers did not fold into helices. Yet, this result does not necessarily mean ether functionalized mPPEs are unable to fold into helices in all circumstances. In Section 3.4, mPPE1, which has ether functional groups, was indeed shown to fold in acetonitrile. Thus, it is possible that higher chain length mPPE3 material may fold into helical structures in moderately polar solvents, but this assertion was not directly tested. Also, the fact that a helical structure was observed for mPPE3 in aqueous solutions indicates that π-stacking is not impossible for the polymer. Thus, while we observed some level of destabilization of the helical structure due to ether substituents, this effect may be mitigated by controlling other factors, such as solvent and polymer chain length, as demonstrated in Section 3.3.

3.5 Summary and conclusions

This chapter presented an experimental and modeling study of two mPPEs having both ester and nitrile functional groups arranged in alternating positions on the mPPE polymer backbone. The folding behaviors of these polymers in acetonitrile and chloroform were evaluated using an REMD simulation procedure, while the actual secondary structures of the mPPEs were characterized using UV absorbance and fluorescence emission spectroscopies. Data from the modeling efforts suggested that reducing the number of ester functional groups on the mPPE would lessen its folding propensity. This effect, however, could be balanced by simply increasing the number of π-stacking aromatic pairs. The modeling predictions agree with the experimental data for the folding behaviors of the synthesized mPPEs. Denaturation titration experiments indicated that the helical structures of these mPPEs were less energetically favored than those of an all ester functionalized mPPE of comparable chain length. These results suggest an approach for the synthesis of helical, functionalized mPPEs and demonstrate the use of modeling as a guide for experimental work.

Following our proposal of new mPPE structures, this study reports the synthesis of two new mPPE copolymers that contain both ester and amine functional groups. These two polymers are not only soluble in water and many common protic solvents, but are also able to fold into helical conformations in water. Thus, these mPPEs could potentially be candidates for new biological applications. The folding behaviors were also reported for these two mPPEs and their precursors, which contain amine protecting groups, in a range

of aprotic and protic polar solvents. Results showed that an *endohelix* methoxy functional group can significantly destabilize the helical structure, but steric interactions between these groups were not a factor. The chain length of the polymers was also found to play a decisive role in determining secondary structure. For example, the longer chain length sample of mPPE4 exhibited helical structures in acetonitrile and several alcohol solvents, while the shorter chain length sample of the same polymer did not fold in these solvents. These results provide further insight to the process of secondary structure formation by mPPEs and related polymers.

3.6 Experimental section

Chromatography methods:

Flash column chromatography employed 200-400 mesh silica gel, 60A from Sigma-Aldrich, with N_2 pressure. Thin layer chromatography used silica gel 60 F_{254} plates from Merck; chemical locations were determined using UV light.

Characterization methods:

NMR spectra were obtained in the Chemistry Department at Clemson University using a 300 MHz Bruker Avance for both ^{1}H and ^{13}C spectra, with either $CDCl_3$ (99.8% atom D, Acros Organics) or DMSO-d_6 (99.9% atom D + 1% v/v TMS, Cambridge Isotope Laboratories). UV/VIS absorption spectra were measured using a Varian Bio 50 UV/Visible Spectrophotometer, with a 1 cm path length quartz cell (Starna Cells, Inc.).

The absorbance was measured from 200 nm to 400 nm, using a 0.5 nm step between measurements. Fluorescence spectra were obtained in the Chemistry Department at Clemson University, using a Photon International - Fluorescence Photometer system, using a quartz cell with a 1 cm path length (Starna Cells, Inc.), an excitation wavelength of 290 nm, an emission scan from 300 nm to 500 nm, and a 1 nm step. All UV absorbance and fluorescence emission spectra were recorded at temperatures ranging from 20 to 25 °C. Low resolution matrix assisted laser desorption ionization (MALDI) experiments were performed on a Bruker Autoflex Matrix-Assisted Laser Desorption/Ionization - Time-of-Flight Mass Spectrometer, using 2,4-dihydroxy benzoic acid (DHB) or 7,7,8,8-tetracyanoquinodimethane (TCNQ) as the matrix. Elemental analysis for Carbon, Hydrogen and Nitrogen were conducted using a Perkin Elmer CHNS/O analyzer 2400, calibrated using acetanilide as a standard. Liquid Chromatography-Electron Ionization Mass Spectrometry was conducted on a modified Extrel (Pittsburgh, PA, USA) Benchmark Thermabeam LC/MS quadrupole mass spectrometer in the Chemistry Department at Clemson University.[46] The LC mobile phase was 50:50 Methanol/milliQwater, with a flow rate of 1 mL/min. The particle beam interface, which delivers the dry analyte particles to the Electron Ionization (EI) source, consists of a nebulizer and a momentum separator. The nebulizer temperature was at 100 °F with a He sheath gas. The momentum separator was heated to 144 °F and the EI source block was heated to 275 °F. The EI filament voltage was 70 eV and detection was done with an electron multiplier at 1400 V. Gel Permeation Chromatography (GPC) was conducted in the Material Science Department at Clemson University using Water Breeze

system equipped with a UV/VIS detector. It was calibrated using narrow molecular weight distribution polystyrene standards (from 400 to 1,000,000 Da). The specific GPC method involved isocratic chloroform flow at 1 mL/min, a UV detector set at 254 nm, and an HR 5E SEC column (range from 2K to $4x10^6$ Da).

All chemicals used in the synthesis procedures in this study were commercially available and used as received. The monomers and polymers were synthesized using as written or modified procedures from previous studies.[14, 32, 33, 47, 48] The yield reported for each reaction is based on the moles of product and the moles of the limiting reactant.

2-(2-(2-methoxyethoxy)ethoxy)ethyl 3,5-diiodobenzoate (1):[47, 48] To a 50 mL round bottom flask (dried then cooled under N_2) was added dichloromethane (25 mL), triethylamine (TEA) (0.78 mL, 3 equiv.). 3,5-diiodobenzoyl chloride (1 g, 2.46 mmol) was then slowly added to the above solution, followed by 2-(2-(2-methoxyethoxy)ethoxy)ethanol (or triethylene glycol monomethyl ether) (TgOH) (0.5 mL, 1.25 equiv.) under continuous stirring at 0 °C (immersed in an ice bath). Next, the ice bath was removed and the solution was allowed to warm to room temperature, then stirred for 24 h. The solution was washed with a 20% NH_4Cl solution (using distilled water) (repeated 2 times), the organic layer was collected and dried with anhydrous Na_2SO_4. Dichloromethane was removed using a rotary vacuum evaporator to collect crude product as a yellow solid. The crude product was purified using flash column chromatography with 1:1 v/v hexane/ethyl acetate as the eluent to collect 0.72 g of product as white crystals (1.38 mmol, yield 56%). TLC $R_f = 0.25$ (1:2 v/v hexane/ethyl

acetate), ^{1}H NMR (300MHz, $CDCl_3$, δ): ^{1}H NMR (300MHz, $CDCl_3$, δ): 8.34 (d, J = 1.6 Hz, Ar H, 2H), 8.24 (t, J = 1.6 Hz, Ar H, 1H), 4.47 (t, J = 4.7 Hz, -COO-C*H*$_2$-CH_2-, 2H), 3.84 (t, J = 4.7 Hz, -COO-CH_2-C*H*$_2$-O-, 2H), 3.65-3.72 (m, -O-C*H*$_2$– C*H*$_2$-O-C*H*$_2$-CH_2-O-CH_3, 6H), 3.54-3.57 (m, –C*H*$_2$-O-CH_3, 2H), 3.39 (s, 3H, -CH_2-OC*H*$_3$); ^{13}C NMR (300MHz, $CDCl_3$, δ): 163.54, 149.16, 137.68, 133.24, 94.27, 71.87, 70.59, 70.53, 68.92, 64.67, 58.95. MALDI DHB matrix (m/z) $[M+H]^+$ calcd for $C_{14}H_{19}O_5I_2^+$ 520.932; found, 521.592; $[M+Na]^+$ calcd for $C_{14}H_{18}O_5I_2Na^+$ 542.914; found, 543.708. LC-MS EI (70 Ev), m/z (relative intensity): 429.95 (2), 399.12 (91), 355.34 (100), 327.55 (49), 260.38 (44), 230.79 (31), 201.81 (31). Anal. Calcd. for $C_{14}H_{18}I_2O_5$ (520.102): C 32.33; H 3.49. Found: C 32.27; H 3.20.

2-(2-(2-methoxyethoxy)ethoxy)ethyl 3,5-diethynylbenzoate (2): To a 100 mL round bottom flask (dried then cooled under N_2, equipped with a magnetic bar) was added 2-(2-(2-methoxyethoxy)ethoxy)ethyl 3,5-diiodobenzoate (chemical (1)) (1.0 g, 1.92 mmol), $Pd_2(dba)_3$ (35.3 mg, 0.02 equiv.), CuI (14.7 mg, 0.04 equiv.), and triphenyl phosphine ($P(Ph)_3$) (101.3 mg, 0.2 equiv.). The flask was sealed, purged with N_2, and evacuated (repeated 3 times) to remove any moisture. Next, dried diisopropyl amine (DIPA) (6.2 mL) dried toluene (61.3 mL), and trimethylsilylacetylene (TMSA) (1.060 mL, 4 equiv.) were added using syringes. The solution was stirred and heated to 78 °C in an oil bath for 48 h. The solution was cooled to room temperature, passed through a short silica gel column, using ethyl acetate as the eluent. Solvent was removed by rotary vacuum evaporator to produce a yellow oil crude product. The crude product was purified via flash column chromatography, using ethyl acetate as the eluent, to collect the final

product as yellow oil. This oil was then dissolved in tetrahydrofuran (50 mL), to this solution was added tetra-*n*-butylammonium fluoride (TBAF 1.0 M in tetrahydrofuran with 5% water) (5.4 mL) and then mixture stirred at room temperature for 15 min. The black solution was passed through a short silica gel column, using ethyl acetate as the eluent, to collect the crude product as an orange oil. The crude product was then purified by flash column chromatography, eluted with 1:2 v/v hexane/ethyl acetate, to collect 0.40 g of a dark yellow solid (1.26 mmol, yield 66%). TLC R_f = 0.3 (1:2 v/v hexane/ethyl acetate). ^{1}H NMR (300MHz, $CDCl_3$, δ): 8.14 (s, Ar H, 2H), 7.77 (s, Ar H, 1H), 4.49 (t, J = 4.7 Hz, -COOC*H*$_2$-CH$_2$-, 2H), 3.85 (t, J = 4.7 Hz, -COOCH$_2$-C*H*$_2$-O-, 2H), 3.65-3.74 (m, -O-C*H*$_2$– C*H*$_2$-O-C*H*$_2$-CH$_2$-O-CH$_3$,6H), 3.54-3.57 (m, –C*H*$_2$-O-CH$_3$, 2H), 3.38 (s, -CH$_2$-OC*H*$_3$, 2H), 3.17 (s, -C≡C*H*, 2H). ^{13}C NMR (300MHz, $CDCl_3$,δ): 164.93, 139.38, 133.32, 130.83, 122.99, 81.79, 78.96, 71.93, 70.70, 70.63, 69.06, 64.58, 59.04. MALDI DHB matrix (m/z) $[M+H]^+$ calcd for $C_{18}H_{21}O_5^+$ 317.139; found, 317.237; $[M+Na]^+$ calcd for $C_{18}H_{20}O_5Na^+$ 339.121; found, 339.297. LC-MS EI (70 Ev), m/z (relative intensity): 227 (7), 212.08 (16), 196.93 (100) 181.96 (6), 153.05 (28), 125.44 (4). Anal. Calcd. for $C_{18}H_{20}O_5$ (316.353): C 68.34; H 6.37. Found: C 67.98; H 6.10.

3,5-diiodo-4-methoxybenzonitrile (3):[14, 32, 33] To a 250 mL round bottom flask with a magnetic stir bar was added dried tetrahydrofuran (THF) (100 mL). The solvent was cooled to 0 °C via immersion in an ice bath. Then, 4-hydroxy-3,5-diiodobenzonitrile (2.0 g, 5.39 mmol) was added to the solvent, followed by methanol (0.26 mL, 1.3 equiv.) and $P(Ph)_3$ (2.119 g, 1.5 equiv.). Diisopropyl azodicarboxylate (DIAD) (1.7 mL, 1.6 equiv.) was slowly added in small portions under continuous stirring to avoid any temperature

excursions above 0 °C. After all chemicals were added, the ice bath was removed and the solution was stirred for 24 h and allowed to warm to room temperature. The solvent was removed using a rotary vacuum evaporator. Then, 50 mL of diethyl ether was added to the flask, and the mixture was stirred at room temperature overnight. The resulting mixture was filtered and a slightly yellow filtrate solution isolated. Diethyl ether was removed from the filtrate solution using a vacuum evaporator to yield a crude pale yellow solid product. The yellow solid was dissolved in dichloromethane and purified via flash column chromatography using 5:1 v/v hexane/dichloromethane as the eluent, yielding 1.6 g of white crystalline product (4.16 mmol, yield 77%). TLC(diethyl ether) $R_f = 0.7$. ^{1}H NMR (300MHz, $CDCl_3$, δ): 8.10 (s, Ar H, 2H), 3.95 (s, Ar-OCH_3, 3H); ^{13}C NMR (300MHz, $CDCl_3$, δ): 163.09, 143.11, 115.31, 111.65, 90.77, 60.95. MALDI TCNQ matrix (m/z) or M^+ calcd for $C_8H_5ONI_2^+$ 384.846; found, 385.600. LC-MS EI (70 Ev), m/z (relative intensity): 383.22 (73), 386.31 (33), 340.43 (7), 258.24 (6), 242.57 (100), 227.84 (15), 214.8 (8), 127.18 (2). Anal. Calcd. for $C_8H_6I_2NO$ (384.950): C 24.96; H 1.57; N 3.63. Found: C 25.35; H 1.24; N 3.69.

Polymerization procedure:[14, 32, 33] To a 25 mL round bottom flask (dried then cooled under N_2, equipped with a magnetic stir bar) was added 3,5-diiodo-4-methoxybenzonitrile (chemical (3), 150 mg, 0.39 mmol) (for mPPE1) or 3,5-diiodobenzonitrile (for mPPE2), 2-(2-(2-methoxyethoxy)ethoxy)ethyl 3,5-diethynylbenzoate (chemical (2), 123.9 mg, 1.0 equiv.), $Pd_2(dba)_3$ (7.2 mg, 0.02 equiv.), CuI (3.0 mg, 0.04 equiv.), and $P(Ph)_3$ (20.4 mg, 0.2 equiv.). The flask was sealed, purged with N_2 and evacuated (three times) to remove any moisture. Then, dried diisopropyl

amine (DIPA) (1.25 mL) and dried toluene (12.4 mL) were added using syringes. The solution was heated to 78 °C and stirred in an oil bath for 24 h. During that time, the initial clear solution turned cloudy. The final solution was cooled to room temperature, and poured into 200 mL of diethyl ether, yielding a yellow precipitate. The dark yellow solid was collected via filtration, re-dissolved in dichloromethane and passed through a short silica gel column, eluting with dichloromethane. The solvent was removed from the yellow solution using a rotator evaporator, and the solid product later was re-precipitated in diethyl ether and further purified by passing through a short silica gel column, eluted with 9:1 v/v chloroform/isopropanol. The polymer was isolated as a light yellow solid. Final recovery of mPPE1 and mPPE2 were 73.9 mg and 78.3 mg, respectively.

mPPE1: ^{1}H NMR (300MHz, $CDCl_3$, δ): 8.24 (br s, Ar H), 7.90 (br s, Ar H, 1H), 7.81 (br s, Ar H), 4.54-4.57 (br m, -COOC*H*$_2$-CH$_2$-), 4.35 (br s, Ar-OC*H*$_3$), 3.88-3.91 (br m , -COOCH$_2$-C*H*$_2$-O-, 2H), 3.66-3.74 (br m, -O-C*H*$_2$–C*H*$_2$-O-C*H*$_2$-CH$_2$-O-CH$_3$), 3.55-3.57 (br m,-C*H*$_2$-O-CH$_3$), 3.37 (br s, -CH$_2$-OC*H*$_3$). GPC result: M_w = 15705 Da, M_n = 8113 Da, M_p = 8970 Da, Polydispersity (Pd) = 1.72. A small amount of very high molecular weight polymer was also detected (less than 0.5% of the total peak area was beyond the range of the polystyrene standards).

mPPE2: ^{1}H NMR (300MHz, $CDCl_3$, δ)**:** 8.22 (br s, Ar H), 7.92 (br s, Ar H) 7.88 (br, Ar H), 7.79 (br, Ar H), 4.55 (br s, -COOC*H*$_2$-), 3.89 (br s, -C*H*$_2$-O-), 3.67-3.75 (br m, -O-C*H*$_2$–C*H*$_2$-O-C*H*$_2$-CH$_2$-O-CH$_3$), 3.54-3.57 (br m, -C*H*$_2$-O-CH$_3$), 3.37 (br s, -CH$_2$OC*H*$_3$). GPC results: M_w = 33926 Da, M_n = 9261 Da, M_p=17929 Da, Pd = 3.66. There was a

relatively large amount of very high molecular weight polymer detected (accounted for 12.9% of the total peak area and had a weight higher than that of the polystyrene standards).

2-(2-(2-methoxyethoxy)ethoxy)ethyl 3,5-bis(phenylethynyl)benzoate (4 or mPPEa): To a 25 mL round bottom flask (dried then cooled under N_2, equipped with a magnetic stir bar) was added 2-(2-(2-methoxyethoxy)ethoxy)ethyl 3,5-diiodobenzoate (chemical (1), 200 mg, 0.38 mmol), ethynylbenzene (117.2 mg, 3.0 equiv.), $Pd_2(dba)_3$ (9.6 mg, 0.02 equiv.), CuI (4.0 mg, 0.04 equiv.), and ($P(Ph)_3$) (27.2 mg, 0.2 equiv.). The flask was sealed, purged with N_2 then evacuated (3 times) to remove any moisture. Then, dried diisopropyl amine (1.67 mL) and dried toluene (16.5 mL) were added using syringes. The solution was stirred at 78 °C in an oil bath for 24 h. The solution was later cooled to room temperature and passed through a short silica gel column (using ethyl acetate as the eluent). Solvent was removed using a rotary vacuum evaporator to collect crude product as a yellow oil. The crude product was later purified by flash column chromatography, eluted with 1:2 v/v hexane/ethyl acetate, to collect the final product as yellow oil. TLC (1:2 v/v hexane/ethyl acetate) $R_f = 0.63$. ^{1}H NMR (300MHz, $CDCl_3$, δ): 8.18 (d, J = 1.4 Hz, Ar H, 2H), 7.88 (m, J = 1.4 Hz, Ar H, 2H), 7.55-7.59 (m, Ar H, 4H), 7.38-7.40 (m, Ar H, 6H), 4.54 (t, J = 4.7 Hz, -COOC*H*$_2$-, 2H), 3.89 (t, J = 4.7 Hz, -COOCH$_2$-C*H*$_2$-O-, 2H), 3.67-3.75 (m, -O-C*H*$_2$–C*H*$_2$-O-C*H*$_2$-CH$_2$-O-CH$_3$, 6H), 3.54-3.57 (m, –C*H*$_2$-O-CH$_3$, 2H), 3.38 (s, -OC*H*$_3$, 3H).^{13}C NMR (300MHz, $CDCl_3$, δ): 138.5, 132.2, 131.8, 130.9, 128.8, 128.5, 124.1, 90.9, 87.6, 71.93, 70.7, 70.6, 69.1, 64.6, 59.0. GPC results: M_w = 623 Da, M_n = 688 Da, M_p= 667 Da, Pd = 1.1. MALDI DHB matrix (m/z) $[M+H]^+$ calcd

for $C_{30}H_{29}O_5^+$ 469.201; found, 469.741; $[M+Na]^+$ calcd for $C_{30}H_{28}O_5Na^+$ 491.183; found, 491.839. LC-MS EI (70 Ev), m/z (relative intensity): 408.32 (4), 364.6 (18), 347.7 (67), 320.9 (88), 304 (100), 290.22 (14), 275.1 (95). Anal. Calcd. for $C_{30}H_{28}O_5$ (468.548): C 76.90; H 6.02. Found: C 74.82; H 5.80.

(6) and (7):[14, 32, 33] To a 500 mL round bottom flask (equiped with a magnetic stirring bar, immersed in an ice bath) was added dried tetrahydrofuran (THF, 100 mL) and BH_3·THF 1M (100 mL). 3,5-diiodo-4-methoxybenzonitrile (2.20 g, 5.72 mmol) (chemical (3), for mPPE1) or 3,5-diiodobenzonitrile (2.20 g, 6.20 mmol, for mPPE2) was dissolved in 20 mL THF, then added into the BH_3 solution in small portions in order to ensure that the temperature did not rise above 0 °C. Several boiling stones were added to the flask, a condenser was attached and solution was refluxed for 24 h. Then, the solution was cooled to room temperature and methanol (30 mL) was slowly added to completely deactivate any excess BH_3 (hydrogen bubbles were expected). The solvent was removed from the resulting solution by rotary evaporation; the remaining solid was wash with methanol (20 mL) then evaporated (repeated 3 times). Methanol (20 mL) was added into the flask to form a milky mixture. The flask was cooled to 0 °C in an ice bath, and HCl gas was slowly bubbled through the solution for 15 min. Later, diethyl ether (50 mL) was added and HCl was bubbled for approximately 15 min or until the solution became clear. Additional diethyl ether (250 mL) was added, and a slightly brown solid (amine salt) precipitated from the solution. The solid was filtered, washed with diethyl ether (10 mL) (repeated 3 times), and dried under vacuum to collect the amine salt for mPPE3 (2.10 g,

5.10 mmol, yield 89%) or the amine salt for mPPE4 (1.53 g, 3.87 mmol, yield 62%). Note that the amine salt must be protected immediately to avoid degradation.

***tert*-butyl 3,5-diiodophenyl-4-methoxybenzylcarbamate (6)**:[14, 32, 33] To a 25 mL round bottom flash was added the amine salt (2.10 g, 5.10 mmol) from the reduction procedure, $(Boc)_2CO$ (1.50 g, 1.4 equiv.), N,N-dimethylformamide (DMF) (10 mL) and solid NaOH (0.38 g, ~ 2.0 equiv.). The mixture was stirred for 15 min; then, distilled water (2 mL) was added in one portion. The flask was covered with aluminum foil then stirred overnight at room temperature. After 24 h, distilled water (10 mL) was added, then the mixture was filtered and washed with excess distilled water to obtain the protected amine product as a white solid (1.75 g, 3.58 mmol, yield 70%). ^{1}H NMR (300MHz, $CDCl_3$, δ):7.68 (s, Ar H, 2H), 4.88 (br s, NH), 4.21 (d, J = 5.4 Hz, -C*H*$_2$-NH-, 2H), 3.85 (s, Ar-OC*H*$_3$, -3H,), 1.45 (s, C-(C*H*$_3$)$_3$, 9H); ^{13}C NMR (300MHz, $CDCl_3$, δ): 157.5, 156.1, 138.6, 91.4, 60.7, 43, 28.4. MALDI DHB matrix (m/z) $[M+Na]^+$ calcd for $C_{13}H_{17}I_2NO_3Na^+$ 511.920; found, 512.530. LC-MS EI (70 Ev), m/z (relative intensity): 430 (61), 386.2 (24), 371.3 (36), 359.36 (4), 261.41 (100), 246.55 (43), 230.75 (23), 202.78 (9). Anal. Calcd. for $C_{13}H_{17}I_2NO_3$ (489.091): C 31.93; H 3.50; N 2.86. Found: C 32.88; H 3.83; N 2.84.

***tert*-butyl 3,5-diiodophenylbenzylcarbamate (7)**: To a 25 mL round bottom flash was added the amine salt (1.53 g, 3.87 mmol) from the reduction procedure, $(Boc)_2CO$ (1.15 g, 1.4 equiv.), N,N-dimethylformamide (DMF) (10 mL) and solid NaOH (0.31 g, ~ 2 equiv.). The mixture was stirred for 15 minutes; then, distilled water (5 mL) was added in

one portion. The mixture became homogeneous for a short time, but later turned to milky solution. The flask was covered with aluminum foil, then stirred overnight at room temperature. After 24 h, distilled water (10 mL) was added, and the mixture was filtered to collect a crude product as white solid product. The product was purified via flash column chromatography, where it was eluted first with hexane then ethyl acetate. The product containing fractions (confirmed using TLC) were vaporized to collect the final protected amine as a white solid (0.92 g, M = 459.037 Da, 2.00 mmol, yield 52%). TLC (hexane: ethyl acetate 1:1 (v/v)) $R_f = 0.48$. ^{1}H NMR 300MHz, $CDCl_3$, δ): 7.96 (t, J = 1.5 Hz, Ar H, 1H), 7.59-7.60 (m, Ar H, 2H), 4.90 (br s, -NH-), 4.22 (d, J = 5.5 Hz, -C*H*$_2$-NH-, 2H), 1.48 (s, -C(C*H*$_3$)$_3$, 9H); ^{13}C NMR (300MHz, $CDCl_3$, δ):146.75, 143.98, 143.20, 135.60, 94.9, 43.2, 28.36. MALDI DHB matrix (m/z) $[M+Na]^+$ calcd for $C_{12}H_{15}I_2NO_2Na^+$ 481.909; found, 483.496. LC-MS EI (70 Ev), m/z (relative intensity): 401.18 (53), 383.28 (1.8), 356.33 (17), 341/43 (20), 275.18 (27), 257.39 (11), 231/68 (100), 215.92 (15). Anal. Calcd. for $C_{12}H_{15}I_2NO_2$ (459.065): C 31.40; H 3.29; N 3.05. Found: C 32.25; H 3.25; N 2.95.

p_mPPE3 and p_mPPE4:[14, 32, 33] To a 25 mL round bottom flask (dried then cooled under N_2, equipped with a magnetic bar) was added (chemical (6),for p_mPPE3) or (chemical (7) for p_mPPE4) (100 mg), (2) (66.5 mg, 1.0 equiv.), $Pd_2(dba)_3$ (3.8 mg, 0.02 equiv.), CuI (1.6 mg, 0.04 equiv.), and $P(Ph)_3$ (10.6 mg, 0.2 equiv.). The flask was sealed, purged with N_2 and evacuated (three times) to remove any moisture. Then, dried diisopropyl amine (0.8 mL) and dried toluene (8.0 mL) were added using syringes. The solution was heated to 78 °C and stirred in an oil bath for 24 h. The final dark brownish

yellow solution was filtered through a short silica gel column to remove the catalyst, eluted with 9/1 v/v chloroform/iso-propanol to collect a clear, yellow solution containing the polymer and monomers. This solution was concentrated and the polymer was separated from the monomer by flash column chromatography, where it was eluted with 1:2 v/v hexane/ethyl acetate, then with 9:1 v/v chloroform/isopropanol. The polymer containing fractions were concentrated, precipitated in hexane then filtered, washed with hexane to collect the final polymer as a pale yellow solid. Final recovery of p_mPPE3 and p_mPPE4 were 94.5 mg and 112 mg, respectively.

p_mPPE3 ^{1}H NMR (300MHz, $CDCl_3$, δ): 8.19 (br s, Ar H), 7.88 (br s, Ar H**),** 7.47 (br s, Ar H, 2H), 5.01 (br s, -NH), 4.53-4.56 (br m, -COOCH_2), 4.23 (br s, -CH_2-NH-), 4.19 (br s, -OCH_3), 3.87-3.90 (br m, $COOCH_2$-CH_2-O-, 2H), 3.65-3.74 (br m, -O-CH_2–CH_2-O-CH_2-CH_2-O-CH_3), 3.55-3.58 (br m, -CH_2-O-CH_3), 3.37 (br s, CH_2-O-CH_3), 1.49 (br s, -C(CH_3)$_3$). GPC result: M_w = 21735 Da, M_n = 9622 Da, M_p = 9481 Da, Polydispersity = 2.26. Small amounts of very high molecular weight polymer were also detected (they were beyond the range of the polystyrene standards).

p_mPPE4 ^{1}H NMR (300MHz, $CDCl_3$, δ): 8.13-8.15 (br m, Ar H), 7.80-7.85(br m, Ar H), 7.63 (br s, Ar H), 7.47 (br s, Ar H), 5.10 (br s, -NH), 4.52 (br s, -COOCH_2), 4.35 (br s, -CH_2-NH-), 3.88 (br s, -COO-CH_2-CH_2-O-), 3.66-3.74 (m, -O-CH_2–CH_2-O-CH_2-CH_2-O-CH_3), 3.53-3.56 (m, -CH_2-O-CH_3), 3.36 (s, -CH_2-OCH_3), 1.51 (s, -C(CH_3)$_3$). GPC results: for long chain length fraction M_w = 12591 Da, M_n = 6822 Da, M_p = 8661 Da, Pd = 1.85; for short chain length fraction: M_w = 8376 Da, M_n = 3818 Da, M_p = 3812 Da, Pd

= 2.19. Small amounts of very high molecular weight polymer were also detected (they were beyond the range of the polystyrene standards).

mPPE3 and mPPE4:[14, 32, 33] To a 25 mL round bottom flask (dried then cooled under N_2, equiped with a magnetic stirring bar and immersed in an ice bath) was added HCl 4M/ Dioxane (2 mL). Later, p_mPPE3 (or p_mPPE4) in (10 mg in 3 mL chloroform solutions) was slowly added. The mixture was stirred at 0 °C for 1 h, during that time, the clear yellow solution turned to a milky solution, and yellow precipitate was observed. The yellow suspension was filtered; the solid was washed on the filter paper with diethyl ether (3 times, 50 mL each). The solid was dissolved in methanol (10 mL), concentrated then re-precipitated by adding diethyl ether (50 mL). The resulting mixture was filtered and washed with diethyl ether to collect the final product as a pale yellow solid (7.5 mg).

mPPE3 1H NMR (300MHz, DMSO-d6, δ): 8.54 (br s, -NH_3+), 8.02-8.08 (br m, Ar H), 7.79-7.87 (br m, Ar H), 4.45 (br s, -COOCH_2-), 4.18 (br s, -OCH_3), 4.04 (br s, -CH_2-NH-), 3.78-3.97 (br m, -CH_2-O-CH_2-), 3.50-3.60 (br m, -CH_2-O-CH_2-), 3.17 (s, -OCH_3).

mPPE4 1H NMR (300MHz, DMSO-d6, δ): 8.62 (br s, -NH_3+), 7.83-8.26 (br m, Ar H), 4.43 (br s, -COOCH_2-), 4.08-4.12 (br m, -CH_2-NH-), 3.77 (br s, -CH_2-O-CH_2-), 3.50-3.62(br m, -CH_2-O-CH_2-), 3.18 (s, -OCH_3).

3.7 References

1. Hill, D. J.; Mio, M. J.; Prince, R. B.; Hughes, T. S.; Moore, J. S., *Chem. Rev.* **2001,** *101* (12), 3893-4011.

2. Nelson, J. C.; Saven, J. G.; Moore, J. S.; Wolynes, P. G., *Science* **1997,** *277* (5333), 1793-1796.

3. Ray, C. R.; Moore, J. S., Supramolecular Organization of Foldable Phenylene Ethynylene Oligomers. In *Poly(Arylene Ethynylene)S: from Synthesis to Application*, **2005**; Vol. 177, pp 91-149.

4. Lahiri, S.; Thompson, J. L.; Moore, J. S., *Journal of the American Chemical Society* **2000,** *122* (46), 11315-11319.

5. Nguyen, H. H.; McAliley, J. H.; Batson, W. A.; Bruce, D. A., *Macromolecules* **2010,** *43* (14), 5932-5942.

6. Cary, J. M.; Moore, J. S., *Org. Lett.* **2002,** *4* (26), 4663-4666.

7. Nguyen, H. H.; McAliley, J. H.; Bruce, D. A., *Macromolecules* **2011,** *44* (1), 60-67.

8. Prince, R. B.; Saven, J. G.; Wolynes, P. G.; Moore, J. S., *J. Am. Chem. Soc.* **1999,** *121* (13), 3114-3121.

9. Stone, M. T.; Heemstra, J. M.; Moore, J. S., *Acc. Chem. Res.* **2006,** *39* (1), 11-20.

10. Hill, D. J.; Moore, J. S., *Proc. Natl. Acad. Sci. U.S.A.* **2002,** *99* (8), 5053-5057.

11. *High performance computing*. http:/citi.clemson.edu/hpc (accessed on 08/03/2011).

12. Adisa, B.; Bruce, D. A., *J. Phys. Chem. B* **2005,** *109* (15), 7548-7556.

13. Adisa, B.; Bruce, D. A., *J. Phys. Chem. B* **2005,** *109* (42), 19952-19959.

14. Adisa, B. Synthesis, characterization and modeling of novel polymers and polymer-templated mesoporous materials. Ph.D. Dissertation, Clemson University, Clemson, SC, U.S.A., **2005**.

15. Bekker, H.; Dijkstra, E. J.; Berendsen, H. J. C., *Supercomputer* **1993,** *10* (2), 4-10.

16. Berendsen, H. J. C.; van der Spoel, D.; Vandrunen, R., *Comput. Phys. Commun.* **1995,** *91* (1-3), 43-56.

17. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J., *J. Comput. Chem.* **1997,** *18* (12), 1463-1472.

18. Hukushima, K.; Nemoto, K., *J. Phys. Soc. Jpn.* **1996,** *65* (6), 1604-1608.

19. Lindahl, E.; Hess, B.; van der Spoel, D., *J. Mol. Model.* **2001,** *7* (8), 306-317.

20. van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C., *J. Comput. Chem.* **2005,** *26* (16), 1701-1718.

21. van der Spoel, D.; Lindahl, E.; Hess, B.; van Buuren, A. R.; Apol, E.; Meulenhoff, P. J.; Tieleman, D. P.; Sijbers, A. L. T. M.; Feenstra, K. A.; van Drunen, R.; Berendsen, H. J. C. GROMACS USER MANUAL version 4.5.4. www.gromacs.org/documentation (accessed 08/03/2011).

22. *Materials Studio*, 4.4; Accelrys Software Inc: San Diago, CA, 2008.

23. Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R., *J. Chem. Phys.* **1984,** *81* (8), 3684-3690.

24. Parrinello, M.; Rahman, A., *J. Appl. Phys.* **1981,** *52* (12), 7182-7190.

25. Jorgensen, W. L.; Madura, J. D.; Swenson, C. J., *J. Am. Chem. Soc.* **1984,** *106* (22), 6638-6646.

26. Jorgensen, W. L., *J. Phys. Chem.* **1986,** *90* (7), 1276-1284.

27. Jorgensen, W. L.; Maxwell, D. S.; TiradoRives, J., *J. Am. Chem. Soc.* **1996,** *118* (45), 11225-11236.

28. Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L., *J. Phys. Chem. B* **2001,** *105* (28), 6474-6487.

29. Price, M. L. P.; Ostrovsky, D.; Jorgensen, W. L., *J. Comput. Chem.* **2001,** *22* (13), 1340-1352.

30. Rizzo, R. C.; Jorgensen, W. L., *J. Am. Chem. Soc.* **1999,** *121* (20), 4827-4836.

31. Watkins, E. K.; Jorgensen, W. L., *J. Phys. Chem. A* **2001,** *105* (16), 4118-4125.

32. Arnt, L.; Tew, G. N., *J. Am. Chem. Soc.* **2002,** *124* (26), 7664-7665.

33. Arnt, L.; Tew, G. N., *Macromolecules* **2004,** *37* (4), 1283-1288.

34. Huang, Y. Q.; Fan, Q. L.; Liu, X. F.; Fu, N. N.; Huang, W., *Langmuir* **2010,** *26* (24), 19120-8.

35. Bahary, W. S.; Jilani, M., *J. Appl. Polym. Sci.* **1993,** *48* (9), 1531-1538.

36. Boni, K. A.; Sliemers, F. A.; Stickney, P. B., *J. Polym. Sci. A: Polym. Chem.* **1968,** *6* (9pa2), 1579-1591.

37. Castignolles, P.; Gaborieau, M., *J. Sep. Sci.* **2010,** *33* (22), 3564-3570.

38. Radke, W.; Simon, P. F. W.; Muller, A. H. E., *Macromolecules* **1996,** *29* (14), 4926-4930.

39. Shultz, A. R.; Mccullou.Cr; Bridgman, A. L.; Hadsell, E. M., *J. Polym. Sci. A: Polym. Chem.* **1972,** *10* (2), 273-282.

40. Huang, S. L.; Tour, J. M., *J. Org. Chem.* **1999,** *64* (24), 8898-8906.

41. Bunz, U. H. F., The ADIMET Reaction: Synthesis and Properties of Poly(dialkylparaphenylenetynylene)s. In *Modern arene chemistry*, Astruc, D., Ed. Wiley-VCH: Weinheim, **2002**; pp 225-229.

42. Lim, K. C.; Fincher, C. R.; Heeger, A. J., *Phys. Rev. Lett.* **1983,** *50* (24), 1934-1937.

43. Li, C. J.; Slaven, W. T.; Chen, Y. P.; John, V. T.; Rachakonda, S. H., *Chemical Communications* **1998,** (13), 1351-1352.

44. Stone, M. T.; Moore, J. S., *Org. Lett.* **2004,** *6* (4), 469-472.

45. Tan, C. Y.; Pinto, M. R.; Kose, M. E.; Ghiviriga, I.; Schanze, K. S., *Adv. Mater.* **2004,** *16* (14), 1208-1212.

46. Castro, J.; Pregibon, T.; Chumanov, K.; Marcus, R. K., *Talanta* **2010,** *82* (5), 1687-1695.

47. Khan, A.; Hecht, S., *Chem. Commun.* **2004,** (3), 300-301.

48. Khan, A. U. H. Phenylene ethynylene foldamers: from synthesis to tubular scaffolding and photoswitchable helices. Ph.D. Dissertation, Freien Universität Berlin, Berlin, Germany, **2005**.

CHAPTER 4

SYNTHESIS OF MANGANESE(SALEN) COMPLEXES USING IMINE FUNCTIONALIZED *meta*-POLYPHENYLENE ETHYNYLENE HELICAL STRUCTURES

4.1 Introduction

meta-Poly(phenylene ethynylene) (mPPE) materials have attracted much attention due to their ability to form biomimetic, highly ordered helical conformations.[1-3] Their open helical structures, formed with six aromatic rings per turn, are stabilized by π-stacking interactions between pairs of overlapping aromatic rings at the average distance of 0.4 nm. The formation of these helical structures also leads to the ordering of the mPPE functional groups on the exterior and interior walls of the helical polymer. This ability of the mPPE polymer to self-organize into helical conformations affords one the possibility of selecting mPPE functional groups whose ordered proximity can be favorably exploited. For example, the three dimensional ordering of polymer functional groups can be used to i) further stabilize the helical polymer conformation (via the formation of intra- and inter-turn hydrogen or covalent bonds),[4-6] ii) prohibit the transport of species through the core of the columnar polymer, or iii) provide selective binding sites for charged or molecular species.

There have been several experimental studies that have already exploited the functional group ordering that arises from mPPE polymers forming helical secondary structures. For example, Hecht and Khan[5] proposed forming semi-rigid mPPE organic nanotubes via the formation of covalent bonds between overlapping pendants. Prince et al.[7] employed the coordination of nitrile functional groups inside the helical cavity to form a silver metal complex. Cary and Moore[4] utilized the adjacent amide functional groups to form hydrogen bonds, which further stabilized the mPPE helical structures. Finally, Nguyen et al.[6] demonstrated in their modeling work (see Chapter 5), that hydrogen bonds could be formed between overlapping functional groups inside the helix cavity, and these hydrogen bonds significantly increase the stability of helical conformations.

In previous chapters, we presented a class of mPPE materials having an alternating arrangement of ester and select other functional groups on the polymer back bone.[8-9] This alternating pattern of functional groups permitted the mPPEs to be functionalized with a variety of chemical groups without compromising the folding ability or solubility of the polymer. Specifically, the presence of the ester functional groups stabilized overlapping aromatics rings by increasing the inter-turn π-stacking interactions between aromatic groups, while the long ether pendants attached to the ester functional groups enhanced the polymer solubility in a variety of solvents.[2] The present work builds upon those developments.

In this study, we demonstrate that select reactions with appropriately oriented (overlapping) functional groups on the mPPE helical structures can lead to many novel

applications for these materials. Of particular interest is the synthesis of an mPPE structure that could be used as an ordered polymer support for catalytically active metal complexes. If successful, these materials would be the first example of a homogeneous catalyst being supported on a structurally uniform non-biological polymer support. These catalysts could have the advantageous property of being made soluble during reactions via the proper choice of solvents (i.e., properties of a homogeneous catalyst), then the polymer can be precipitated and easily removed from the reactive solution (i.e., properties of a heterogeneous catalyst) by slightly changing the solvent conditions. Our initial efforts in this area focused on the synthesis of an mPPE polymer with ester and alkyl imine functional groups arranged in an alternating pattern. The flexibility of the alkyl (ethyl) linkage between the mPPE aromatic rings and the imine functional groups provide sufficient flexibility for the imines groups to rearrange and favorably bind to select transition metals, such as manganese and cobalt, forming salen type metal complexes (see Figure 4.1). The homogeneous variants of these salen complexes have been shown to be highly active and selective catalysts for epoxidation and epoxide ring opening reactions.[10-14] Though other researchers have attempted to anchor these complexes to solid supports (e.g., silica, alumina, and random oriented polymers), invariably their efforts met with mixed results because each of these supports has very irregular surface features, which leads to each metal complex being in a slightly different environment that in turn effects their selectivity and activity.[15-19] In summary, the successful anchoring of metal complexes to structurally regular polymers may lead to a new approach for the synthesis of polymer supported, heterogeneous catalysts, which

interestingly mimics the strategy employed with many biological catalysts, such as enzymes.[20]

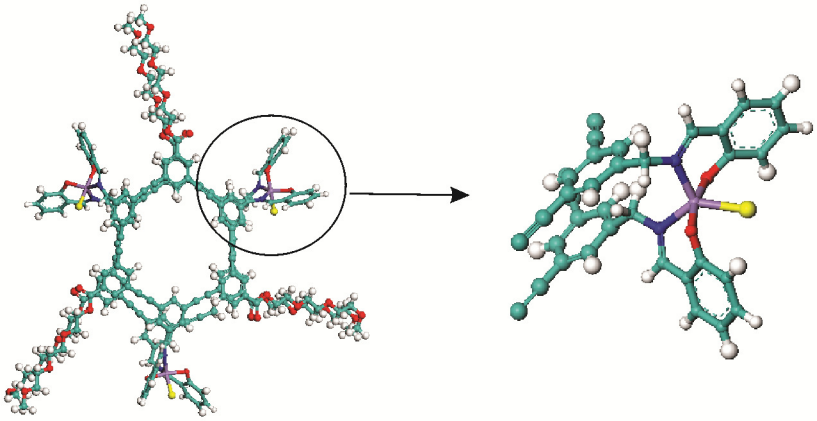


Figure 4.1 Molecular rendering of the manganese(salen) complex (Mn(salen)-mPPE complex). Atom colors for carbon, chlorine, hydrogen, manganese and oxygen are teal, yellow, white, purple and red, respectively.

An additional benefit of the salen complex is that it locks together stacked polymer units in a way that secures the helical conformation of the mPPEs, thereby, providing an alternative method to synthesize mPPE organic nantubes. For example, in the approach of Hecht and Khan[5], covalent C-C bonds were formed between the pendants of two overlapping aromatic rings. However, the length of the C-C bond is only approximately 0.15 nm, whereas the average distance between two overlapping aromatic rings is approximately 0.4 nm.[8, 21] As a result, the presence of these short C-C bonds could lead to structural distortions of the helical polymer. The flexible linkages between the imine functional groups and the mPPE backbone eliminate the rigid constraint of the C-C bonds

seen with earlier efforts, and thus better preserve the arrangement of aromatic groups that are observed with functionalized mPPE polymer.

In this chapter, we present the synthesis and characterization of a manganese(salen) complex (Mn(salen)-mPPE complex) that employ imine functionalized mPPE helical structures as the metal ligands. The material characterization efforts included reactions studies that examined the catalytic activity of the supported complex for epoxidation reactions. In addition, we propose two approaches for the synthesis of chiral variants of this complex for future studies examining the enantioselectivity of these supported catalysts.

4.2 Synthesis of the Mn(salen)-mPPE complex

The imine mPPE was synthesized from the amine functionalized mPPE presented in Chapter 3. The synthesis route of the imine functionalized mPPE and the Mn(salen)-mPPE complex are presented in Figure 4.2 and Figure 4.3, respectively. The chain length of the protected amine mPPE was estimated using GPC results that employed polystyrene with narrow molecular weight distribution standard and an mPPE analog with three aromatic ring as calibration standards. Previous results in Chapter 3 indicate that the synthesized mPPE chains contain a minimum of 14 aromatic rings, which is a sufficient chain length form stable mPPE helical conformations under suitable solvent conditions (as evidenced by the spectroscopy data presented in Figure 4.5).

The imine mPPE formation reaction was conducted in a 1:1 v/v methanol/chloroform mixture. A K_2CO_3 solution in a 75:25 v/v methanol/water mixture was added to the solution at regular intervals to inhibit the precipitation of the amine intermediate. The formation of the Mn(salen)-mPPE complex was carried out in a 9:1 v/v methanol/chloroform mixture at room temperature. These conditions promote mPPE helical conformations, allowing the imine functional groups to align in such a way as to form the salen ligands for the complex. The polarity of the solution was also sufficient for the inorganic precursors to be soluble. A long reaction time was expected for complex formation because the reaction was conducted at room temperature (as compared to the reflux conditions (70-80 °C) normally used to form the complex.[11, 13-14] This lower temperature was used to avoid the destabilizing effect of elevated temperature on the mPPE helical structures.[2-3]

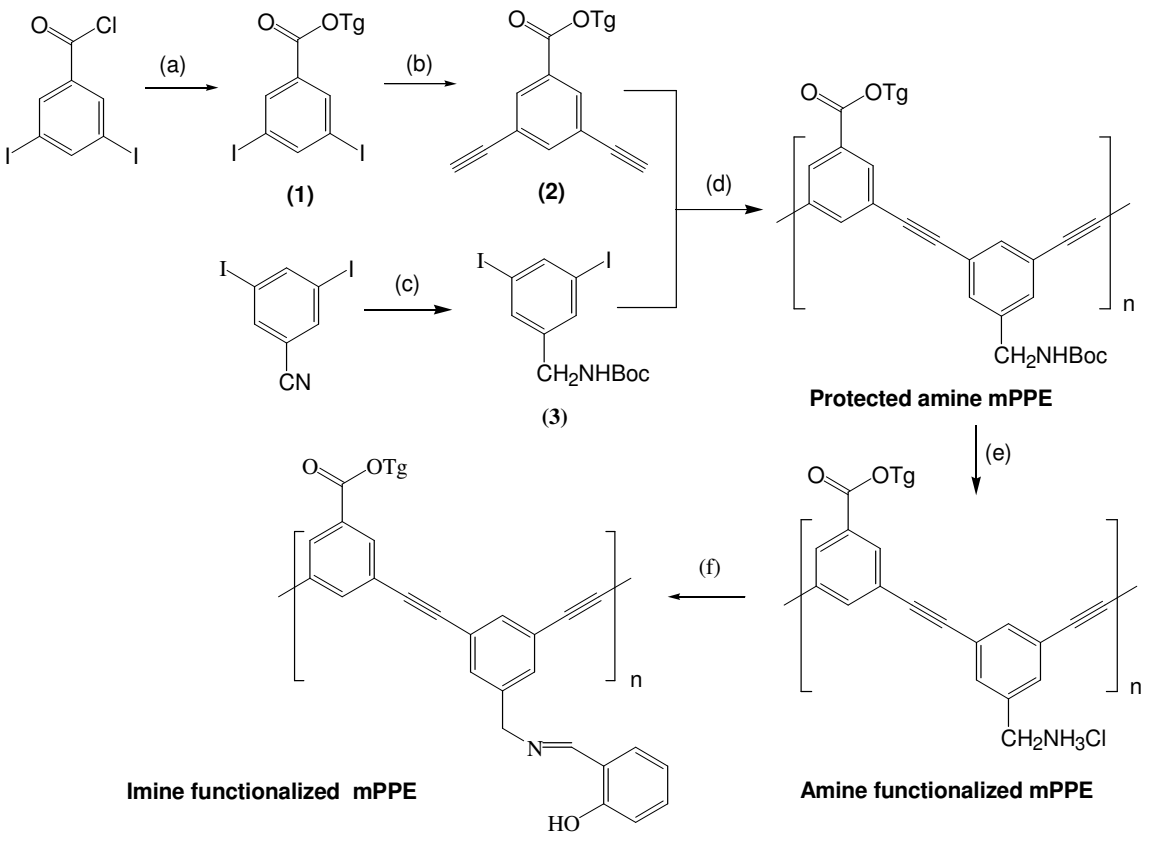


Figure 4.2 Synthesis route for the imine functionalized mPPE. Reagents: (a) 2-(2-(2-methoxyethoxy)ethoxy)ethanol (TgOH), triethylamine (TEA), CH_2Cl_2, r.t., 24 h, 56%; (b) trimethylsilylacetylene (TMSA), $Pd_2(dba)_3$, CuI, $P(Ph)_3$, diisopropyl amine (DIPA), toluene, 78 °C, 24 h; then tetra-*n*-butylammonium fluoride (TBAF), THF, 2h, 66%; (c) BH_3·THF, THF, reflux, 24 h; then $(Boc)_2CO$, N,N-dimethylformamide (DMF), NaOH, H_2O, 24 h, 52%; (d) $Pd_2(dba)_3$, CuI, $P(Ph)_3$, DIPA, toluene, 78 °C, 24 h; (e) HCl 4M/dioxane, CH_2Cl_2, 0 °C, 2 h; (f) 2-hydroxybenzaldehyde, K_2CO_3, $CHCl_3$:CH_3OH (1:1 v/v), reflux, 2 h.

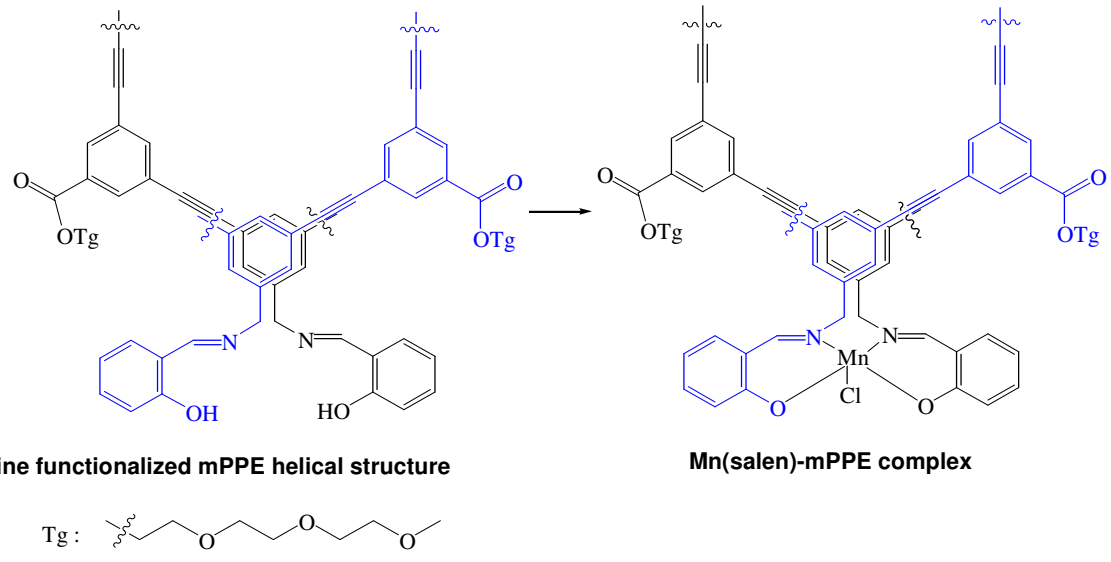


Figure 4.3 Synthesis route for the Mn(salen)-mPPE complex. The overlapping aromatic rings (blue and black colors) are to depict two consecutive layers of the mPPE helical structure; the salen complex is formed between these two overlapping imine functional groups. Reagents: $Mn(OAc)_2 \cdot 4H_2O$, $CHCl_3$:CH_3OH (1:9 v/v), r.t., 5 day; then LiCl, MeOH, air, 24 h.

4.3 Mn(salen)-mPPE complex characterization

The Mn(salen)-mPPE complex was characterized by several analytical techniques, including FT-IR (see Figure 4.4).[13, 22-29] The presence of the metallo salen complex absorption band at 1540 cm^{-1} and the shift of the free C=N stretching vibration band from 1631 cm^{-1} to 1609 cm^{-1} support the formation of a Mn(salen) complex (see spectra in

Figure 4.4a and Figure 4.4b). These features are similar to the variations in FT-IR spectra observed with Jacobsen's homogeneous variant of the Mn(salen) catalyst and its salen ligand precursor, shown in Figure 4.4c and Figure 4.4d, and are strong indicators of the formation of the Mn(salen)-mPPE complex. However, the broad band spanning from 1631 cm^{-1} to 1609 cm^{-1} shows an incomplete shift, indicating there were free imine functional groups in the final complex product. That is, not all imine functional groups on the mPPE helical structure participated in the formation of metal complexes. This is an expected result that follows naturally from the fact that the imine functional groups must rearrange and react in pairs to form the salen complex, leaving some imine groups excluded by their already paired neighbors.

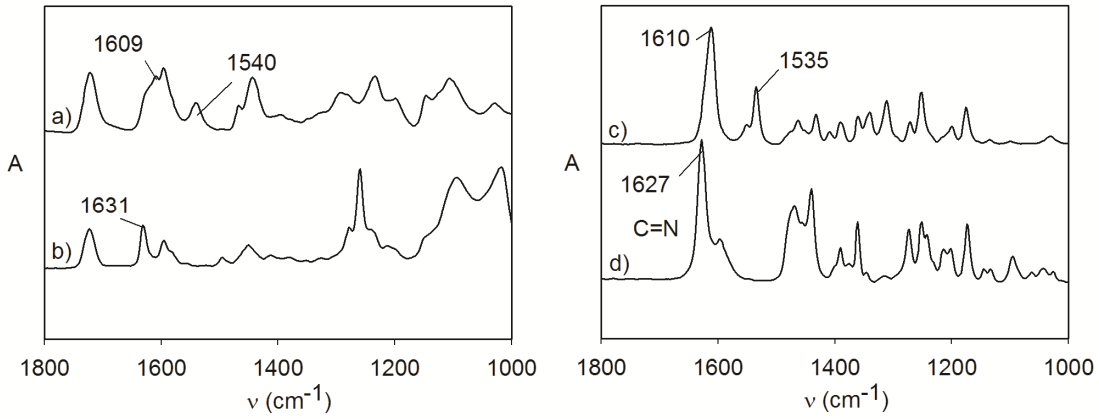


Figure 4.4 Absorption FT-IR spectra of a) Mn(salen)-mPPE complex, b) imine functionalized mPPE, c) Jacobsen's homogeneous Mn(salen) catalyst, and d) Jacobsen's precursor salen ligand.

To quantify the extent of complexation that had occurred, the manganese loading of the imine functionalized mPPEs was determined using a spectrophotometric method,[30] showing the Mn containing polymer had 1.07 wt% of manganese, indicating a low reaction yield. If all of the imine functional groups had formed Mn(salen) complexes then the calculated loading of manganese would have been ~5 wt %; thus, approximately 20% of the imine functional groups present on the helical polymer formed Mn(salen) complexes.

The folding behavior of the imine functionalized mPPE can be determined from the spectroscopic data shown in Figure 4.5. The decrease in UV absorbance intensity at 306 nm for the mPPE in acetonitrile as compared to that in chloroform indicates that the formation of the helical conformation is favored in acetonitrile, while the disordered state is favored in chloroform.[1-3] Whereas, the UV spectrum of the Mn(salen)-mPPE complex in chloroform is almost identical to that in acetonitrile, providing evidence that after the metal complex formed, the mPPE helical structures were stable in both acetonitrile and chloroform – clearly indicating that the formed metal complexes have prevented the helical polymer from unfolding in chloroform. The typical absorption peaks of the salen ligand and the salen complex are not visible because of the strong absorbance of the polymer backbone in the 320-280 nm region. The FT-IR and UV absorption spectra provide clear evidence that the Mn complexes was formed between imine functional groups located at neighboring inter-turn positions on the helical mPPE support (see Figure 4.1).

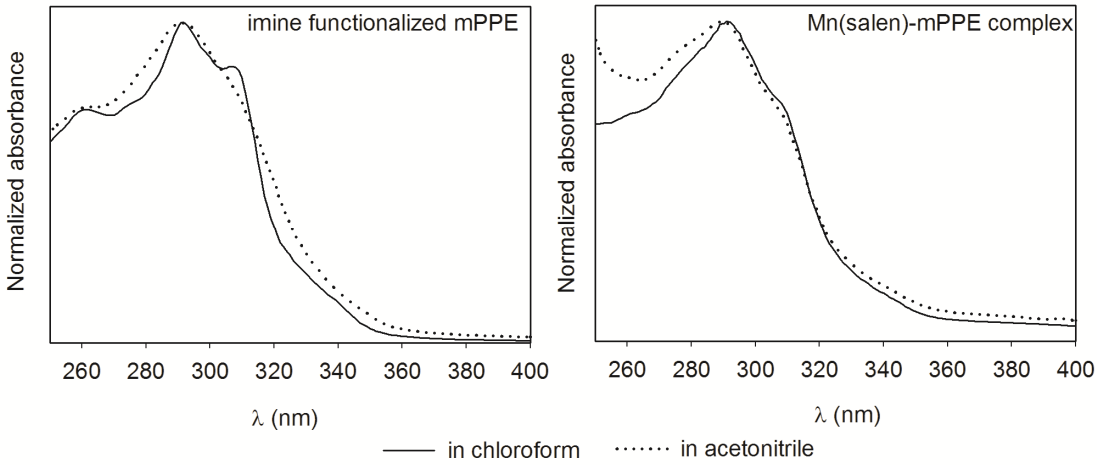


Figure 4.5 UV absorbance spectra of the imine functionalized mPPE (left) and the Mn(salen)-mPPE complex (right) in chloroform (solid line) and acetonitrile (dotted line).

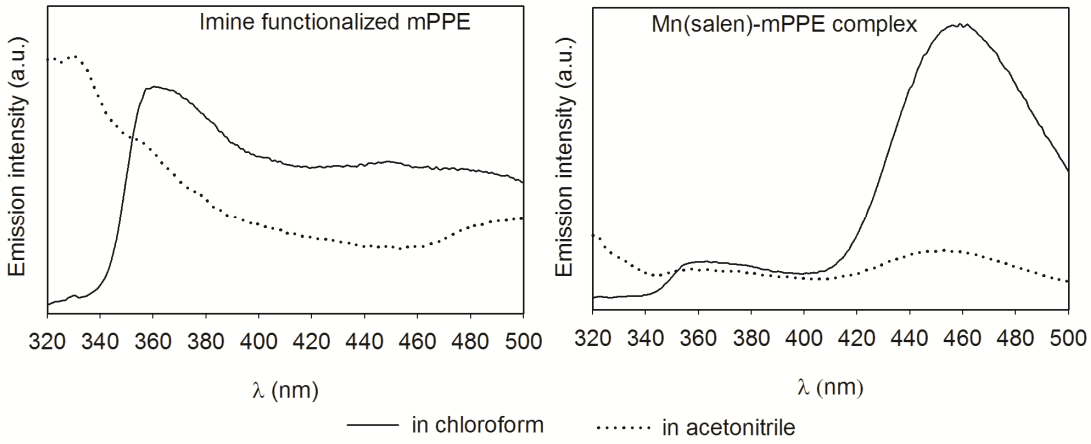


Figure 4.6 Fluorescence spectra of imine functionalized mPPE (left) in chloroform (solid line) and in acetonitrile (dotted line), and of the Mn(salen)-mPPE complex (right) in chloroform (solid line) and in acetonitrile (dotted line).

Additional evidence for the existence of the mPPE helical conformation in the Mn(salen)-mPPE complex can be derived from the characteristic fluorescence emission spectra in chloroform and in acetonitrile, as in previous studies.[2-3, 25, 31-32] If as predicted, the Mn(salen) contains ligands from neighboring inter-turn functional groups on the mPPE, then no mPPE structural transition should occur when changing the solvent from acetonitrile to chloroform; therefore, the Mn(salen)-mPPE complex emission spectra are expected to be similar in both studied solvents (see Figure 4.6). For the imine functionalized mPPE, the absence of the monomer emission peak at 360 nm, due to the excimer effect, indicates the formation of overlapping aromatic pair in acetonitrile. This is a common indicator for the formation of helical structures in solution, as described in Chapter 3 and in previous studies.[2-3, 25, 31-32] Similar absences of the monomer peak are also observed in emission spectra of the Mn(salen)-mPPE complex, both in chloroform and in acetonitrile. Likewise, the fluorescence emission spectra corresponding to the monomer peak are absent from the emission spectra of the Mn(salen)-mPPE complex. A strong emission peak exists at 460 nm for the complex in chloroform, while in acetonitrile, a similar but weaker peak is observed at that wavelength. However, the reason this emission peak is not observed in the emission spectra of the imine functionalized mPPE in acetonitrile is not clear. It could result from the self-quenching of the free oriented salen rings or because of the formation of salen ring excimer pairs, which could further be stabilized by hydrogen bonds formed between their hydroxyl groups. On the mPPE helical structure in acetonitrile the imine functional groups are hypothesized to arrange such that they form overlapping aromatic pairs, thus, creating

another excimer effect and quenching the emission peak at 460 nm. On the other hand, the imine functional groups on the Mn(salen)-mPPE complex are effectively locked by covalent bonds, preventing the formation of excimer pairs. The likely two arrangements of imine functional groups on the imine functionalized mPPE helical structure and on the Mn(salen)-mPPE complex are shown in Figure 4.7.

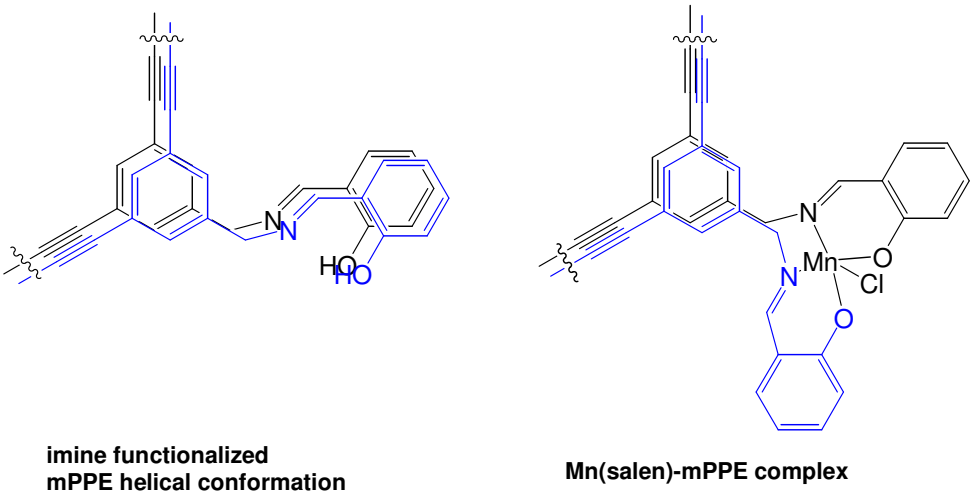


Figure 4.7 The predicted arrangements of imine functional groups on the imine functionalized mPPE in its helical conformation and on the Mn(salen)-mPPE complex. The blue and black colors depict two consecutive helix layers. The structure on the left shows the possible excimer pair by the rearrangement of the two overlapping imine functional groups on the mPPE helical structure.

4.4 Catalytic testing of the Mn(salen)-mPPE complex for styrene epoxidation reactions

Based upon prior catalytic testing by others of a variety of supported Mn(salen) complexes,[10-11, 13-14, 33] the synthesized Mn(salen)-mPPE complex is expected to have some catalytic activity for epoxidation reactions. Therefore, the epoxidation kinetics of styrene were determined using sodium hypochlorite (NaOCl) as the oxidant. The reaction conditions are similar to those published earlier.[10-11, 13-14] The polymer backbone was found to be stable under these reaction conditions for an extended period of time, with or without the presence of the metal complex. This is evident from preliminary testing using diphenyl acetylene (the dimer analog of the mPPE having two aromatic rings) in the exact same conditions used for the styrene epoxidation reaction (see Figure 4.8). After 48 h, there was no observed changed in the diphenyl acetylene concentration nor were any degradation product observed.

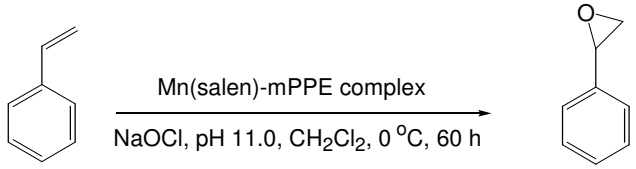


Figure 4.8 Styrene epoxidation reaction using Mn(salen)-mPPE complex as catalyst.

During the reaction using the synthesized Mn(salen)-mPPE complex as catalyst, the epoxide product (styrene oxide) was detected after 12 hours. After 60 hours, analysis using GC-FID showed that the highest observed overall conversion of styrene was 50%, calculated using *n*-decane as the internal standard, but the styrene oxide yield was

observed to be only 2%. Several byproducts were detected using GC-MS (as shown in Figure 4.9).

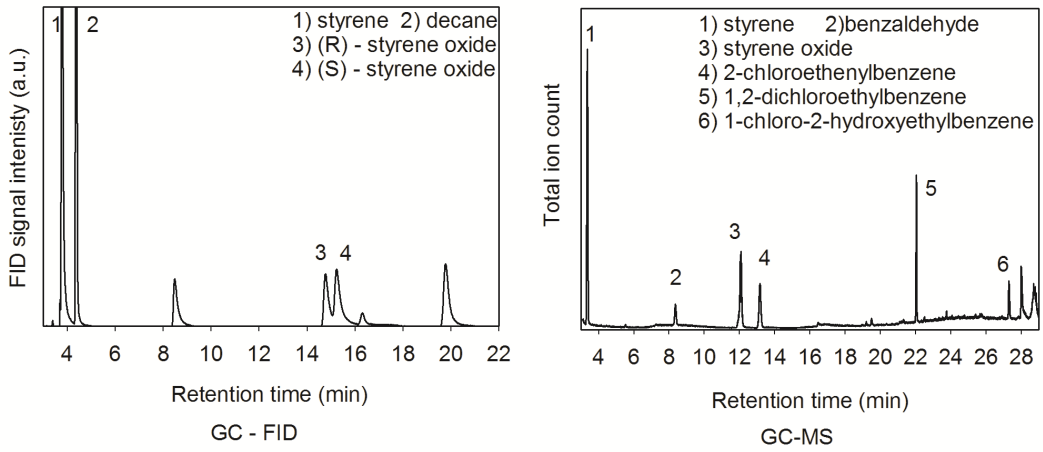


Figure 4.9 GC analysis of samples from the styrene epoxidation reaction using the Mn(salen)-mPPE complex as catalyst. Chemicals were identified using appropriate standards for GC-FID or by searching MS spectra libraries for GC-MS.

These observed byproducts were likely formed directly from styrene through oxidative cleavage routes (benzaldehyde) or via halogenation reactions with the chlorine that was present in the reaction media.[21, 33] The low conversion, yield, and the presence of significant numbers of byproducts can be attributed to several factors. For example, the selected oxidant system, the low catalyst to substrate ratio (i.e. the molar ratio between the Mn(salen) complex and styrene), and the long reaction times which could all lead to the formation of multiple products. The possible hydrolysis reaction of the ester functional groups on the mPPE backbone in strong basic environment could change the solubility of the Mn(salen)-mPPE complex in the organic phase. This could prevent styrene from accessing the active manganese sites, reducing the overall conversion of the

reaction. Further, there could be deleterious distortions of the Mn(salen)-mPPE complex caused by the conformational strain imposed by the mPPE helical structures, which could severely limit catalyst activity and selectivity.

Continuing experimental work in our group is focused on i) the optimization and scaled up of the Mn(salen)-mPPE complex forming reaction; ii) additional characterization and validation of the supported catalyst structure; and iii) further evaluation of the material's catalytic activity for styrene epoxidation using different oxidant systems other than NaOCl.

4.5 Preliminary results for the synthesis of chiral Mn(salen)-mPPE complexes

The result from the preliminary catalyst testing of the achiral Mn(salen)-mPPE catalyst, though somewhat uninspiring, did show that the material exhibited catalytic activity but with low chemoselectivity and yield to the desired epoxide product. More importantly, these efforts provide evidence that the polymer backbone is stable under the epoxidation reaction using NaOCl as the oxidant. Thus, further improvement of this concept is merited. One of the important aspects of a catalyst for olefin epoxidation reactions is the enatioselectivity, which thus far has not been incorporated into the synthesized Mn(salen)-mPPE complexes. In this section, we present experimental approach for creating enantioselective variants of these catalysts by imparting the chirality of mPPE helical structures to the catalyst. The central idea is to bias the polymer to form only one

type of chiral helical structure, either (+) or (-), prior to forming the Mn(salen) complex. This approach is possible based on the results from earlier studies showing that either a chiral guest[34-35] (a small chiral molecule able to fit into the helix cavity) or chiral pendants[36-38] (side groups containing a chiral center attached to the polymer backbone) could be employed to guide the formation of chiral secondary structures in mPPEs. Figure 4.10 presents the general schemes for these two approaches.

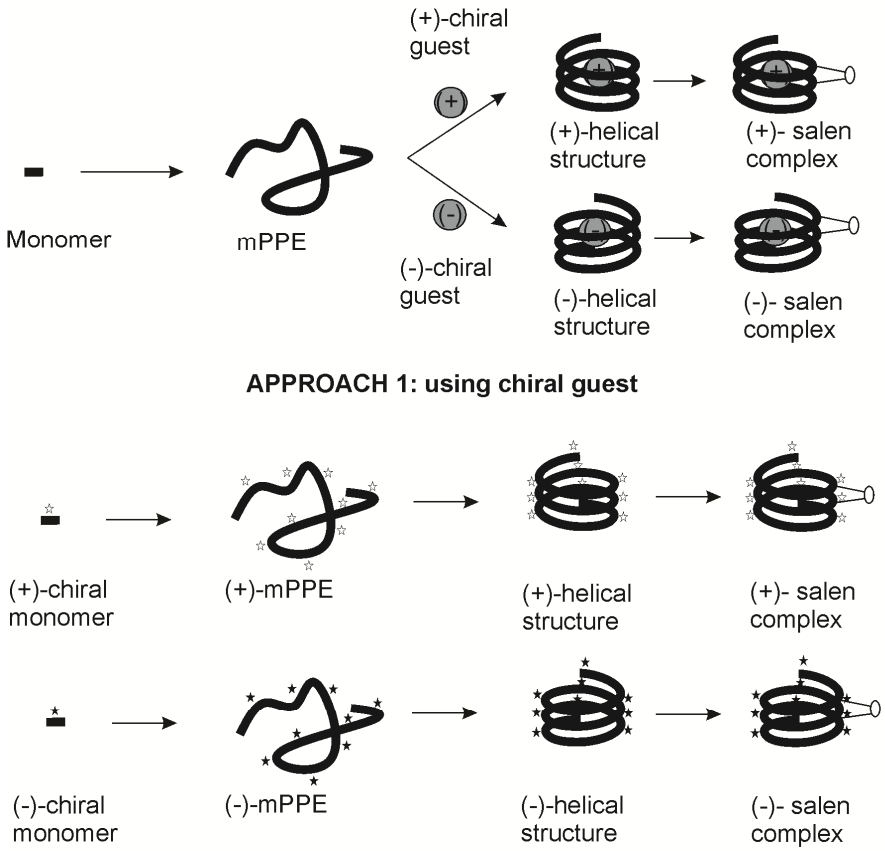


Figure 4.10 General illustration of two proposed approaches to synthesize chiral Mn(salen)-mPPE complexes.

In the first approach, the same imine functionalized mPPE as presented in Section 4.2 is employed. When a chiral molecule of suitable size and functionality, such as α-pinene, is added to an mPPE containing solution, the guest-host interactions between it and the polymer lead to the formation of chiral biased helical structures, observable using circular dichroism (CD).[34-35] For example, Figure 4.11 presents the CD spectra from experiments where the synthesized imine functionalized mPPE and α-pinene molecules are both dissolved in the same solution. The adsorption bands observed in the CD spectra correspond to those of the polymer and are not directly attributable to α-pinene species. Evidence from UV absorbance spectra, similar to those shown in Figure 4.5, indicates that mPPEs fold into helica conformations in these solvent mixtures.

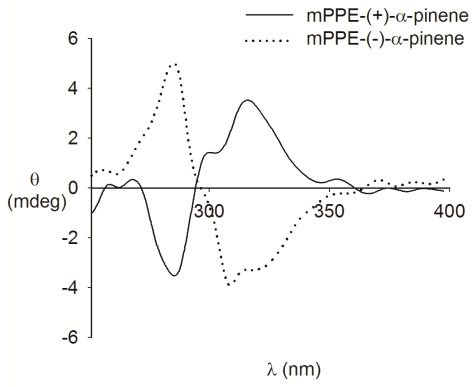


Figure 4.11 CD spectra of the imine functionalized mPPE in the presence of either (-)-α-pinene (dotted line) or (+)-α-pinene (solid line) in an 80:20 v/v acetonitrile/water mixture. Similar spectra were observed for mPPE-α-pinene systems in both methanol/water and in ethanol/water mixed solvent systems. Evidence from UV absorbance spectra indicate that mPPE are in helical conformations.

Preliminary screening for the first approach suggests that the most important aspect for this reaction is determining suitable reaction conditions which could simultaneously promote chiral bias in the folding reaction while allowing formation of the salen complex. Conditions include but are not limited to solvent system, temperature, type of chiral guest and chemical ratio. Experimental results have indicated that to satisfy the first requirement the solvent system must include water, which increases the solvent polarity and forces the non-polar chiral guest to move into the non-polar helical cavity.[34-35, 39] Tanatani reported that this folding bias could take place in an acetonitrile/water mixtures with α-pinene as the chiral guest.[35] However, the salen complex formation has not been observed in that solvent system because of the formation of a manganese (II) acetonitrile complex,[40-41] preventing the slower reaction to form the salen complex. Further, α-pinene is ruled out as a guest because of the potential polymerization or oxidation reaction of either manganese (II) ion or the manganese (II) complex with its double bond.[42-44] We have identified that methanol/water and ethanol/water mixtures are also able to promote chiral bias in the folding reaction. Additionally, results from initial trials with Jacobsen's ligand indicate that manganese(salen) complex could form in 8:2 v/v methanol/water at room temperature. We also found that several saturated chiral molecules, including hydroxyl containing (+)-borneol, (-)-borneol as well as (+)-*cis*-pinane and (-)-*trans*-pinane, could replace α-pinene as the chiral guest.

In order to investigate the second approach in Figure 4.10, we propose the synthesis of two new chiral mPPEs having chiral pendants (shown in Figure 4.12).

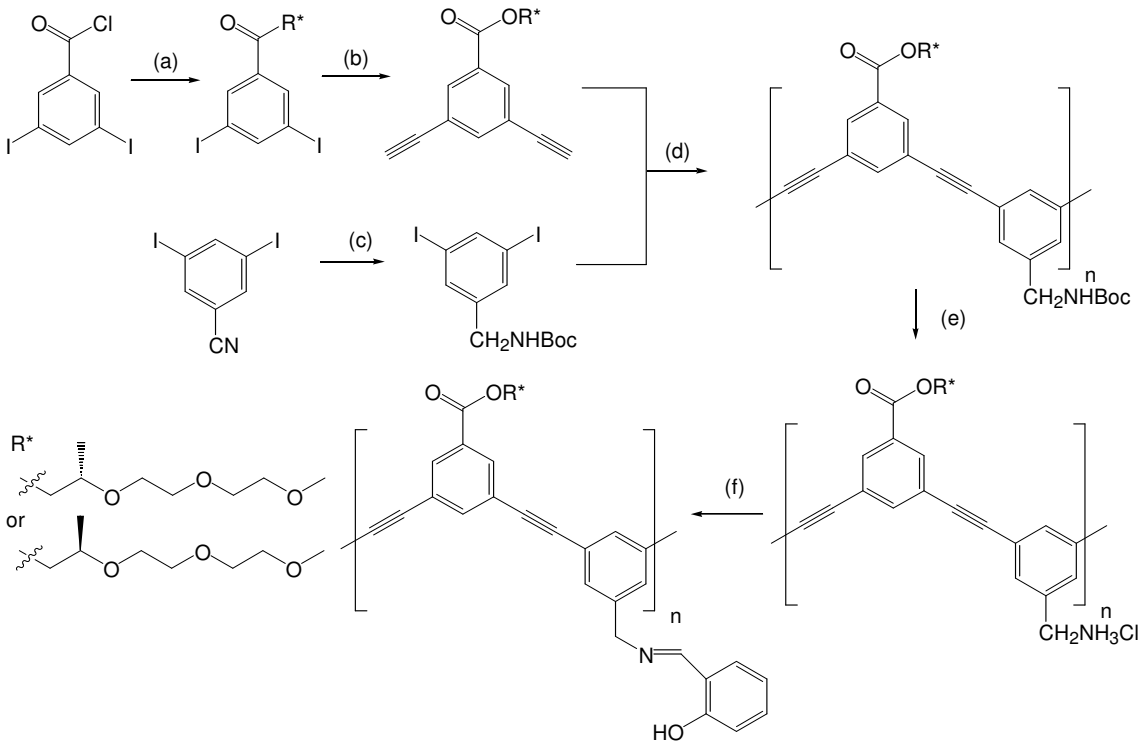


Figure 4.12 The proposed imine functionalized mPPEs with chiral pendants that are to be used in the synthesis of chiral Mn(salen)-mPPE complexes. Reagents: (a) (*R*)-2-(2-(2-methoxyethoxy)ethoxy)propan-1-ol or (*S*)-2-(2-(2-methoxyethoxy)ethoxy)propan-1-ol, TEA, CH_2Cl_2, r.t., 24 h; (b) trimethylsilylacetylene (TMSA), $Pd_2(dba)_3$, CuI, $P(Ph)_3$, diisopropyl amine (DIPA), toluene, 78 °C, 24 h; then *tetra*-n-butylammonium fluoride (TBAF), THF, 2h; (c) BH_3·THF, THF, reflux, 24 h; then $(Boc)_2CO$, N,N-dimethylformamide (DMF), NaOH, H_2O, 24 h; (d) $Pd_2(dba)_3$, CuI, $P(Ph)_3$, DIPA, toluene, 78 °C, 24 h; (e) HCl 4M/dioxane, CH_2Cl_2, 0 °C, 2 h; (f) 2-hydroxybenzaldehyde, K_2CO_3, $CHCl_3$:CH_3OH (1:1 v/v), reflux, 2 h.

These two chiral imine functionalized mPPEs can be synthesized using the procedure presented in Section 4.2, and in a previous study.[9] The ether pendants of the ester functional groups on the achiral monomers are replaced with either the (+) or (-) chiral pendants.[39, 45] The same conditions for the achiral complex, as described in Section 4.2, could be applied to produce the chiral versions of the Mn(salen)-mPPE complex from using these novel chiral polymers.

4.6 Conclusions

This chapter presents the synthesis of a Mn(salen)-mPPE complex using an imine functionalized mPPE helical structure as metal complex ligands. FT-IR spectra provided evidence for metal complex formation, while UV absorbance and fluorescence emission spectra indicated that the manganese complex was formed on mPPE systems in their helical conformation. Preliminary catalytic testing for styrene epoxidation activity using NaOCl as oxidant showed that the Mn(salen)-mPPE complex exhibited moderate catalytic activity, and the polymer backbone was stable under the reaction conditions. From these important results, approaches to synthesize the chiral version of this catalyst were proposed.

4.7 Experimental section

Chromatography methods:

Flash column chromatography employed 200-400 mesh silica gel, 60A from Sigma-Aldrich, with N_2 pressure. Thin layer chromatography used silica gel 60 F_{254} plates from Merck; chemical locations were determined using UV light.

Characterization methods:

NMR spectra were obtained in the Chemistry Department at Clemson University using a 300 MHz Bruker Avance for both ^{1}H and ^{13}C spectra, with either $CDCl_3$ (99.8% atom D, Acros Organics) or DMSO-d_6 (99.9% atom D + 1% v/v TMS, Cambridge Isotope Laboratories). UV/VIS absorption spectra were measured using a Varian Bio 50 UV/Visible Spectrophotometer, with a 1 cm path length quartz cell (Starna Cells, Inc.). The absorbance was measured from 200 nm to 400 nm, using a 0.5 nm step between measurements. Fluorescence spectra were obtained in the Chemistry Department at Clemson University using a Photon International - Fluorescence Photometer system, using a quartz cell with a 1 cm path length (Starna Cells, Inc.), an excitation wavelength of 290 nm, an emission scan from 300 nm to 500 nm, and a 1 nm step. CD experiments were conducted using a Jasco J-810 spectropolarimeter. The CD spectra were recorded in the Chemistry Department at Clemson University from 400 nm to 250 nm as an average of five scans, using a 0.1 nm step size, with a 1 cm path length quartz cell (Starna Cells, Inc.). All UV absorbance, fluorescence emission and CD spectra were recorded at

temperatures ranging from 20 to 25 °C. Permeation Chromatography (GPC) were conducted in the Material Science Department at Clemson University using a Water Breeze system equipped with a UV/VIS detector, which was calibrated using polystyrene with narrow molecular weight distribution standards (from 400 to 1,000,000 Da). The specific GPC method involved isocratic chloroform flow at 1 mL/min, a UV detector set at 254 nm, and an HR 5E SEC column (range from 2K to $4x10^6$ Da). FT-IR spectra were collected in Material Science Department at Clemson University using a Thermo-Nicolet, Magna-TR 550 Spectrometer, equipped with a Thermo Spectro-tech, Foundation series diamond ATR accessary. The IR absorbance spectra were recorded as an average of 16 scans over a wave number range of 640 cm^{-1} to 4000 cm^{-1}.

Conversion and yield calculations for the styrene epoxidation reaction were derived from analysis of products liquids using an HP7890 GC using an HP5 column (30 m × 250 μm × 0.25 μm), helium carrier gas, GC inlet temperature of 220 °C, oven temperature of 50 °C for 2 min, temperature ramp of 10 °C/min to 100 °C, a final hold temperature of 100 °C for 2 min, an inlet gas pressure of 30 psi, a gas feed split ratio of 25:1, an FID detector maintained at temperature of 220 °C, an H_2 flow rate of 40 ml/min, an air inlet pressure of 450 ml/min and helium makeup gas). GC-FID calibration data for styrene, *n*-decane and styrene oxide retention times were 3.36 min, 6.67 min and 8.25 min, respectively (see Figure 4.13).

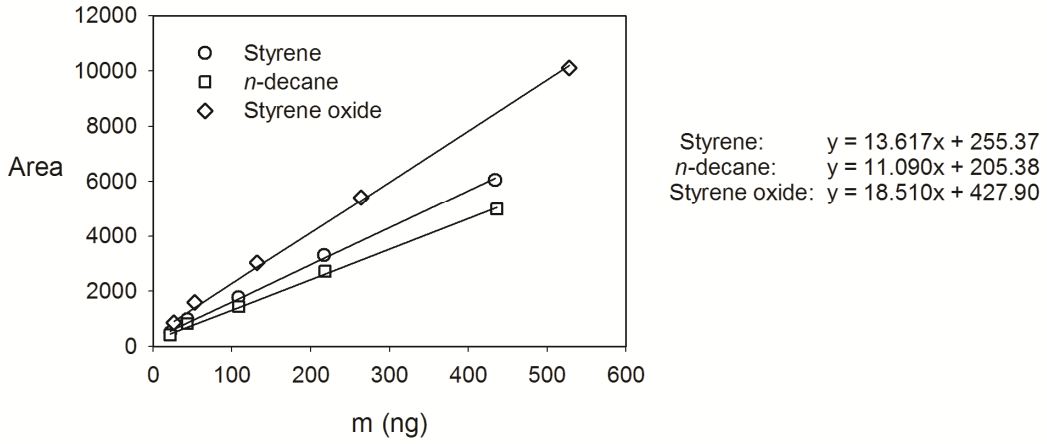


Figure 4.13 GC-FID calibration data for styrene, *n*-decane and styrene oxide.

Data for the calculation of enantiomeric excess for a given styrene epoxidation reaction product was collected via GC using an HP6890 instrument, which included a Restex beta-dex column (30 m × 250 μm × 0.25 μm), helium carrier gas, an inlet temperature of 220 °C, an isothermal oven temperature of 80 °C, a sample gas inlet pressure of 30 psi, a split ratio of 25:1, and an FID detector (220 °C, H_2 at 40 psi, air at 60 psi, and helium makeup gas). The retention times for styrene, *n*-decane, (*R*)-styrene oxide and (*S*)-styrene oxide were 3.76, 4.36, 14.77 and 15.23 min, respectively. Product and byproduct identification for the styrene epoxidation was achieved by GC-MS (HP GC6890-MS5793). Specific machine settings for the analysis included: a JWsc-DBWAXeTR column (30 m × 250 μm × 0.25 μm), helium carrier gas, an inlet temperature of 220 °C, an oven temperature of 110 °C for 15 min, ramped at 10 °C/min to 200 °C, and finally maintained at 200 °C for 10 min, an inlet pressure of 10 psi, a feed gas split ratio of 25:1, and an MS detector (scan mode, solvent delay 3 min, EM voltage 2000V). Compound

identification was achieved via automated comparison of the observed mass spectra to those in the NIST98 MS spectral library.

The yield reported for each reaction is the final recovered yield based on the moles of product and the moles of the limiting reactant. The chemicals (1), (2), (3), the protected amine functionalized mPPE, and the amine functionalized mPPE were synthesized according to the procedure described in Chapter 3. The method used to synthesize the imine functionalized mPPE and the Mn(salen)-mPPE complex were developed from those reported in previous studies.[11, 13, 46]

Imine functionalized mPPE: To a 100 mL round bottom flask was added methanol (10 mL), amine functionalized mPPE (in chloride salt form, 22 mg, ~0.048 mmol of $-NH^{3+}$), chloroform (10 mL) and 2-hydroxybenzaldehyde (20 mg, ~4 equiv.). Several boiling stones were added, a condenser was attached to the flask, and the flask was placed into a heated bath (oil bath). The solution was heated to reflux, and K_2CO_3 (5 mg/mL solution in 75:25 v/v methanol/H_2O, 1.32 mL, ~1 equiv.) was added in five portions (15 min intervals) through the condenser using a micropipette. After all of the salt was added to the mixture, the reaction was refluxed for an additional 1 h. The reaction mixture was then cooled to room temperature and dried under vacuum. Methanol (100 mL) was added, and the mixture was stirred for 5 min then filtered. The solid was dissolved in 5 mL of chloroform then precipitated with large amount of methanol (~100 mL). The precipitate was filtered and washed with methanol (100 mL) to collect the final product as a yellow solid (24.3 mg). The precipitation and filtration steps could be repeated as

needed to remove any remaining 2-hydroxybenzaldehyde. Thin layer chromatography was used to monitor the presence of 2-hydroxybenzaldehyde (TLC, $CHCl_3$) with the polymer $R_f = 0$, 2-hydroxybenzaldehyde $R_f = 0.33$. 1H NMR (300MHz, $CDCl_3$, δ): 8.51(br s, N=C*H*), 8.17 (br s, Ar H), 7.84 (br s, Ar H), 7.67 (br s, Ar H), 7.49 (br s, Ar H), 7.32 (br m, Ar H), 7.02 (br m, Ar H), 6.94 (br m, Ar H), 4.83 (br s, -CH_2-N=C), 4.52 (br s, -COOCH_2-), 3.87(br s, -COOCH$_2$-CH_2-O-,), 3.66-3.72 (br m, -CH_2-CH_2-O-), 3.54 (br s, -CH_2-O-CH$_3$), 3.36 (br s, -OCH_3).

Mn(salen)-mPPE complex: To a 500 mL Erlenmeyer flask was added methanol (270 mL) and chloroform (20 mL). A solution of imine mPPE in chloroform (25.3 mg of imine mPPE in 10 mL chloroform, ~0.048 mmol of imnine functional groups) was then slowly added to the solution, followed by $Mn(OAc)_2 \cdot 4H_2O$ (36 mg, ~1.5 equiv.). The reaction mixture was stirred at room temperature for five days. The resulting light yellow suspension was evaporated until dry. Methanol (50 mL) was then added to form a yellow suspension. The suspension was transferred to a 3-neck 100 mL flask equipped with a condenser, then LiCl (10 mg, ~5 equiv.) was added. Air was bubbled through the suspension for 24 h under constant stirring. The suspension was centrifuged to separate the liquid and the solid phase (2500 rpm, 10 min); the clear liquid phase was slowly removed using a micropipette. The yellow solid was washed with methanol (3 times, 30 mL each), each time the suspension was centrifuged (2500 rpm in 10 min) and the liquid phase was slowly removed. The solid was collected and dried in air for 24 h to collect the final complex as a light yellow powder (14.1 mg).

Styrene epoxidation reaction using Jacobsen's catalyst:[11] To a 25 mL vial was added styrene (208 mg, 2.0 mmol), *n*-decane (200 mg, 1.4 mmol as an internal standard), 4-PPNO (130 mg, 0.38 equiv.), Jacobsen's catalyst (12.7 mg, 1 mol% based on styrene) and CH_2Cl_2 (6 mL). The vial was sealed and placed into an ice bath atop a magnetic stirrer, and the mixture was stirred for 5 min. A 13% solution of NaOCl in water (2.4 mL) was then mixed with a pH 11.00 buffer solution (4.8 mL), and the combined solution was slowly added to the reaction mixture under constant (rapid) stirring. At various time intervals, the stirring was stopped, and the reaction mixture was allowed to stand for 1 min, so as to allow for a complete phase separation. A small sample of the lower layer (50 μL) was then collected, and diluted with CH_2Cl_2 (1 mL). The suspension was filtered through a short silica gel column before being injected into the GC for composition analysis. Complete styrene conversion was achieved after approximately 3 h and the enantiomeric excess was approximately 40%. Thin layer chromatography (TLC, 1:1 v/v hexane/CH_2Cl_2) was also used to follow the reaction, with the styrene $R_f = 0.67$ and the styrene oxide $R_f = 0.33$.

Stability of alkyne groups: The chemical stability of the alkyne moieties under epoxidation reaction conditions was ascertained using the same reaction conditions as those used for the styrene epoxidation reaction, which were presented in the previous section, with and without Jacobsen's catalyst. For these reactions studies, styrene was replaced by diphenyl acetylene (1,2-diphenylethyne) (200 mg, 1.95 mmol). The suspension was filtered through a short silica gel column before being analyzed using a GC equipped with a flame ionization detector (FID). After 48 h, there was no change in

the diphenyl acetylene concentration and no products were detected. TLC (1:1 v/v hexane/CH_2Cl_2) was also used to follow the reaction, where the diphenyl acetylene R_f = 0.61.

Styrene epoxidation reaction using Mn(salen)-mPPE complex as a catalyst: To a 10 mL vial was added styrene (13 mg, 0.125 mmol), *n*-decane (15 mg, 0.106 mmol, as an internal standard), 4-PPNO (7 mg, 0.33 equiv.), Mn(salen)-mPPE complex (10 mg) and CH_2Cl_2 (0.375 mL). The vial was sealed and placed into an ice bath atop a magnetic stirrer, and the solution was stirred for 5 min. A 13% solution of NaOCl in water (0.15 mL) was then mixed with a pH 11.00 buffer solution (0.3 mL), and the combined solution was slowly added to the reaction mixture under constant (rapid) stirring. At various time intervals, stirring was stopped, the reaction mixture was allowed to stand for 1 min for complete phase separation. A small sample of the lower layer (50 μL) was the collected and diluted with CH_2Cl_2 (1 mL). The suspension was filtered through a short silica gel column before being injected into the GC for composition analysis. The observed styrene conversion (using decane as the internal standard) was 49% after 48 h and 51% after 60 h. The relative amounts of styrene oxide enantiomer reaction products were (*R*)-styrene oxide (45%) and (*S*)-styrene oxide (55%). Other reaction products were identified via GC electron ionization mass spectrometry (GC-MS EI (70 eV): styrene (3.35 min), m/z (relative intensity): 104.15 (100), 78.15 (33), 63.25 (4), 51.25(10); benzaldehyde (8.45 min) m/z (relative intensity): 105.95 (100), 77.45 (45), 51.25 (16); styrene oxide (12.00 min) m/z (relative intensity): 119.85 (69), 104 (5), 91 (100), 77.35 (7), 63.25 (14), 51.25 (9); 2-chloroethenylbenzene (13.25 min) m/z (relative intensity): 138.25 (92), 103.25

(100), 77.25 (25), 63.15 (5), 51.15 (12); 1,2-dichloroethylbenzene (22.04 min) m/z (relative intensity): 173.95 (26), 139 (25), 125.25 (100), 103.85 (35), 77.95 (14), 63.25(4), 51 (8); 1-chloro-2-hydroxylethylbenzene (27.3 min) m/z (relative intensity): 156.15 (7), 119.95 (8), 107.35 (100), 79.35 (50), 51.25 (10).

Recovery of Mn(salen)-mPPE complex from epoxidation reaction: After the styrene epoxidation reaction, the reaction mixture was diluted with methanol (10 mL) then centrifuged. The liquid layer was removed, the solid (Mn(salen)-mPPE) was washed with 20 mL methanol then centrifuged (repeated 3 times). The brownish yellow solid was dried in air at room temperature for 24 h to collect the post-reaction supported metal complex (6.1 mg, recovery yield 61%).

Determination of Mn loading in the Mn(salen)-mPPE complex: The Mn loading in Mn(salen)-mPPE samples was determined using spectrophotometric (UV absorption) methods.[30, 47] With this method, all available manganese in the Mn(salen)-mPPE complex is converted to ion form (MnO_4^-):

$$2\ Mn_2^+ + 5IO_4^- + 3H_2O \rightarrow 2MnO_4^- + 5IO_3^- + 6H^+$$

To a 100 mL beaker was added the Mn(salen)-mPPE complex (5 mg) and concentrated nitric acid (5 mL). The mixture was boiled on a heating plate, then H_2O_2 (2 mL) was slowly added under constant stirring using a glass rod until the solution was completely clear. Additional H_2O_2 can be added if necessary. The solution was boiled until dry, then cooled in air to room temperature. After cooling, distilled water (2 mL) was added to the beaker to form a clear solution. The solution was boiled on a heating plate, then H_3PO_4

(0.3 g) was added, followed by KIO_4 (0.05 g). The solution was stirred using a glass rod for 20 min. During this time (after stirring for 5-10 min), the color changed from colorless to deep purple (or pink, depending on the concentration of MnO_4^-). KIO_4 (0.03 g) was added and the solution was boiled and stirred for additional 15 min (during this process if all of the water was evaporated, additional water was added). Then, the solution was allowed to cool to room temperature, and poured into a 25 mL volumetric flask. Distilled water was added until the flask was filled to the specified (to contain) volume. The absorbance of the final solution at 526 nm was determined and the concentration of MnO_4^- was calculated using the calibration data shown in Figure 4.14.

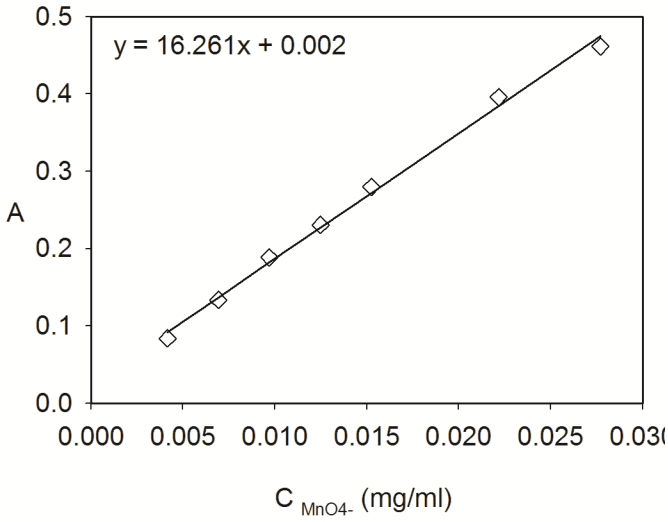


Figure 4.14 UV absorption spectroscopy calibration curve for $KMnO_4$ at 526 nm.

The Mn loading could be calculated from the concentration of MnO_4^- using the following equation:

$$\%\mathrm{Mn}=\frac{\mathrm{C}_{\mathrm{MnO_4^-}}\mathrm{V}_{\mathrm{MnO_4^-}}\mathrm{M}_{\mathrm{Mn}}}{\mathrm{m}_{\mathrm{Mn\text{-}mPPE}}\times\mathrm{M}_{\mathrm{MnO_4^-}}}\times 100\% \qquad (0.1)$$

where *%Mn* equals the weight percentage of Mn in the respective Mn(salen)-mPPE sample; C_{MnO4-} equals the concentration of MnO_4^- in the final test solution (mg/ml); V_{MnO4-} is the volume of the final MnO_4^- test solution; M_{MnO4-} is the ion molecular weight (118.936 g/mol) M_{Mn} is the Mn atomic weight (54.938 g/mol) and $m_{Mn(salen)-mPPE}$ is the weight of Mn(salen)-mPPE in the test sample.

The maximum Mn loading in the Mn(salen)-mPPE complex could be calculated from the following equation, using the ratio of 1 Mn atom per 2 mPPE repeat units:

$$\%\mathrm{max}_{\mathrm{Mn}}=\frac{\mathrm{M}_{\mathrm{Mn}}}{\mathrm{M}_{\mathrm{Mn\text{-}mPPE}}}\times 100\% \qquad (0.2)$$

where $\%max_{Mn}$ is the maximum Mn loading in the Mn(salen)-mPPE complex and $M_{complex}$ is the molecular weight of the Mn(salen)-mPPE complex (using the ratio of 1 Mn atom per 2 mPPE repeat units).

Formation of Jacobsen's catalyst at room temperature: To a 25 mL flask (equipped with a magnetic stirring bar) was added Jacobsen's ligand ((*R,R*)-(−)-*N,N′*-Bis(3,5-di-*tert*-butylsalicylidene)-1,2-cyclohexanediamine) (5 mg), $Mn(OAc)_2{\cdot}4H_2O$ (5 mg, 2 equiv.) and an 8:2 v/v methanol/water mixture (or ethanol/water mixture) (10 mL). The

suspension was stirred at room temperature until a dark brownish red solution (Mn(salen) complex) was observed (after approximately 6 h). The formation of the complex was monitored using TLC (1:4 v/v ethyl acetate/hexane, ligand $R_f = 0.7$, complex $R_f = 0$) and UV absorbance spectroscopy (Figure 4.15).

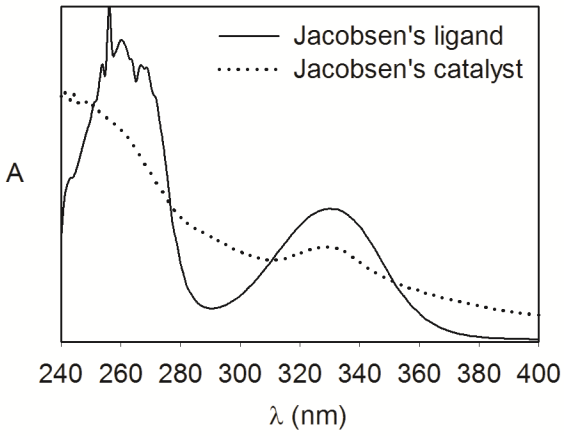


Figure 4.15 UV absorbance spectra of Jacobsen's ligand (solid line) and Jacobsen's catalyst (dotted line).

4.8 References

1. Hill, D. J.; Mio, M. J.; Prince, R. B.; Hughes, T. S.; Moore, J. S., *Chem. Rev.* **2001,** *101* (12), 3893-4011.

2. Nelson, J. C.; Saven, J. G.; Moore, J. S.; Wolynes, P. G., *Science* **1997,** *277* (5333), 1793-1796.

3. Ray, C. R.; Moore, J. S., Supramolecular Organization of Foldable Phenylene Ethynylene Oligomers. In *Poly(Arylene Ethynylene)S: from Synthesis to Application*, **2005**; Vol. 177, pp 91-149.

4. Cary, J. M.; Moore, J. S., *Org. Lett.* **2002,** *4* (26), 4663-4666.

5. Hecht, S.; Khan, A., *Angew. Chem. Int. Ed.* **2003,** *42* (48), 6021-6024.

6. Nguyen, H. H.; McAliley, J. H.; Bruce, D. A., *Macromolecules* **2011,** *44* (1), 60-67.

7. Prince, R. B.; Okada, T.; Moore, J. S., *Angew. Chem. Int. Ed.* **1999,** *38* (1-2), 233-236.

8. Nguyen, H. H.; McAliley, J. H.; Batson, W. A.; Bruce, D. A., *Macromolecules* **2010,** *43* (14), 5932-5942.

9. Nguyen, H. H.; McAliley, J. H.; Bruce, D. A., J. *Polymer Sci. Polymer Chem.* **2011,** *in press*.

10. Adam, W.; Fell, R. T.; MockKnoblauch, C.; SahaMoller, C. R., *Tetrahedron Lett.* **1996,** *37* (36), 6531-6534.

11. Hanson, J., *J. Chem. Educ.* **2001,** *78* (9), 1266-1268.

12. Jacobsen, E. N., *Acc. Chem. Res.* **2000,** *33* (6), 421-431.

13. Zhang, W.; Loebach, J. L.; Wilson, S. R.; Jacobsen, E. N., *J. Am. Chem. Soc.* **1990,** *112* (7), 2801-2803.

14. Zhang, W.; Jacobsen, E. N., *J. Org. Chem.* **1991,** *56* (7), 2296-2298.

15. Choudary, B. M.; Ramani, T.; Maheswaran, H.; Prashant, L.; Ranganath, K. V. S.; Kumar, K. V., *Adv. Synth. Catal.* **2006,** *348* (4-5), 493-498.

16. Xiang, S.; Zhang, Y. L.; Xin, Q.; Li, C., *Chem. Commun.* **2002,** (22), 2696-2697.

17. Yao, X. Q.; Chen, H. L.; Lu, W. R.; Pan, G. Z.; Hu, X. Q.; Zheng, Z., *Tetrahedron Lett.* **2000,** *41* (52), 10267-10271.

18. Zhang, H. D.; Xiang, S.; Li, C., *Chem. Commun.* **2005,** (9), 1209-1211.

19. Zhang, H. D.; Zhang, Y. M.; Li, C., *J. Catal.* **2006,** *238* (2), 369-381.

20. Lehninger, A. L.; Nelson, D. L.; Cox, M. M., *Lehninger principles of biochemistry*. 5th ed.; W.H. Freeman: New York, **2008**; pp 153-162, 183-194.

21. Carey, F. A.; Sundberg, R. J., *Advanced organic chemistry, Part A: Structure and Mechanisms*. 5th ed.; Springer: New York, **2007**; pp 51, 485-497.

22. Baleizao, C.; Gigante, B.; Garcia, H.; Corma, A., *J. Catal.* **2003,** *215* (2), 199-207.

23. Dominguez, I.; Fornes, V.; Sabater, M. J., *J. Catal.* **2004,** *228* (1), 92-99.

24. Frunza, L.; Zgura, I.; Nicolaie, I.; Dittmar, A.; Kosslick, H.; Fricke, R., *J. Optoelectro. Adv. M.* **2005,** *7* (4), 2141-2148.

25. Heinrichs, C.; Holderich, W. F., *Catal. Lett.* **1999,** *58* (2-3), 75-80.

26. Lou, L. L.; Yu, Y.; Yu, K.; Jiang, S.; Dong, Y. L.; Liu, S. X., *Science in China Series B-Chemistry* **2009,** *52* (9), 1417-1422.

27. Tan, R.; Yin, D. H.; Yu, N. Y.; Jin, Y.; Zhao, H. H.; Yin, D. L., *J. Catal.* **2008,** *255* (2), 287-295.

28. Tan, R.; Yin, D. H.; Yu, N. Y.; Zhao, H. H.; Yin, D. L., *J. Catal.* **2009,** *263* (2), 284-291.

29. Ware, D. C.; Mackie, D. S.; Brothers, P. J.; Denny, W. A., *Polyhedron* **1995,** *14* (12), 1641-1646.

30. Palmer, C. www.cas.umt.edu/chemistry/palmer/documents/quantitative_analysis/-Manganese.pdf (accessed 03/23/2010).

31. Birks, J. B., *Chem. Phys. Lett.* **1967,** *1* (7), 304-306.

32. Birks, J. B., *Rep. Prog. Phys.* **1975,** *38* (8), 903-974.

33. Carey, F. A.; Sundberg, R. J., *Advanced organic chemistry, Part B: Reactions and Synthesis*. 5th ed.; Springer: New York, **2007**; pp 1088-1091, 1126-1128.

34. Prince, R. B.; Barnes, S. A.; Moore, J. S., *J. Am. Chem. Soc.* **2000,** *122* (12), 2758-2762.

35. Tanatani, A.; Hughes, T. S.; Moore, J. S., *Angew. Chem. Int. Ed.* **2002,** *42* (2), 325-328.

36. Brunsveld, L.; Prince, R. B.; Meijer, E. W.; Moore, J. S., *Org. Lett.* **2000,** *2* (11), 1525-1528.

37. Prince, R. B.; Brunsveld, L.; Meijer, E. W.; Moore, J. S., *Angew. Chem. Int. Ed.* **2000,** *39* (1), 228-230.

38. Zhao, X. Y.; Schanze, K. S., *Langmuir* **2006,** *22* (10), 4856-4862.

39. Prince, R. B. Phenylene ethynylene foldamers: cooperative conformational transition, twist sense bias, molecular recognition properties and solid-state organization. Ph.D. Dissertation, University of Illinois ar Urbana-Champaign, Urbana-Champaign, IL, U.S.A., **2000**.

40. Gago, S.; Zhang, Y. M.; Santos, A. M.; Kohler, K.; Kuhn, F. E.; Fernandes, J. A.; Pillinger, M.; Valente, A. A.; Santos, T. M.; Ribeiro-Claro, P. J. A.; Goncalves, I. S., *Micropor. Mesopor. Mater.* **2004,** *76* (1-3), 131-136.

41. Sakthivel, A.; Hijazi, A. K.; Yeong, H. Y.; Kohler, K.; Nuyken, O.; Kuhn, F. E., *J. Mater. Chem.* **2005,** *15* (41), 4441-4445.

42. Higashimura, T.; Lu, J.; Kamigaito, M.; Sawamoto, M.; Deng, Y. X., *Macromol. Chem. Physic.* **1993,** *194* (12), 3441-3453.

43. Lu, J.; Liang, H.; Zhang, R. J.; Deng, Y. X., *Acta. Polymerica. Sinica.* **1998,** (6), 698-703.

44. Schuster, C.; Mollmann, E.; Tompos, A.; Holderich, W. F., *Catal. Lett.* **2001,** *74* (1-2), 69-75.

45. Kudo, M.; Hanashima, T.; Muranaka, A.; Sato, H.; Uchiyama, M.; Azumaya, I.; Hirano, T.; Kagechika, H.; Tanatani, A., *J. Org. Chem.* **2009,** *74* (21), 8154-8163.

46. Jacobsen, E. N.; Zhang, W.; Muci, A. R.; Ecker, J. R.; Deng, L., *J. Am. Chem. Soc.* **1991,** *113* (18), 7063-7064.

47. Harris, D. C., *Quantitative chemical analysis*. 7th ed.; W.H. Freeman and Co.: New York, NY, **2007**; pp 357-358.

CHAPTER 5

CONTROLLING THE *meta*-POLY(PHENYLENE ETHYNYLENE) HELICAL CAVITY ENVIRONMENT: HYDROGEN BOND STABILIZED HELICAL STRUCTURES

5.1 Introduction

meta-poly(phenylene ethynylene)s or mPPEs are a class of polymers known for their ability to form helical secondary structures in suitable solvent conditions.[1-3] This special ability not only invites the opportunity for many different applications, but also provides a simple and interesting platform for studying the conformational behavior of macromolecules. Therefore, understanding and controlling the factors that affect their folding behavior is crucial in designing an effective functionalized mPPE for a given application. As the interplay of functional groups on the polymer backbone with themselves and the surrounding solvent can significantly impact the mPPE secondary structure formation, it is especially important to understand the impact that these groups have on that formation.

The properties of a given mPPE are largely determined by the types of functional groups attached at two primary positions on the polymer backbone, as shown in Figure 5.1. Based on their locations on the resulting helical structures, we refer to those in position R_1 as *exohelix* functional groups. These functional groups are located at positions *meta* to the ethynylene linkages and are positioned on the outer wall of the polymer in its helical

conformation. Similarly, we refer to groups at position R_2 as *endohelix* functional groups. These groups are located at positions *ortho* to both ethynylene linkages on a phenyl ring, and they are encased inside the cavity of the mPPE in its helical conformation.

The *exohelix* functional groups of mPPEs have been shown by many experimental and modeling studies to directly affect the solubility and folding behaviors of their respective mPPEs.[2, 4-14] The *endohelix* functional groups, on the other hand, have been employed mostly to control the environment inside the helix cavity. For example, Tanatani et al.[15] employed methyl functional groups to reduce the size of the helix cavity of an mPPE, effectively blocking the entrance of small molecules into the cavity. Additionally, Prince et al.[16] used nitrile *endohelix* functional groups to form a silver ion complex within the helix cavity. The bonds between the ion and the nitrile functional groups of the complex provided an additional stabilizing effect for the helical structure in tetrahydrofuran solution.

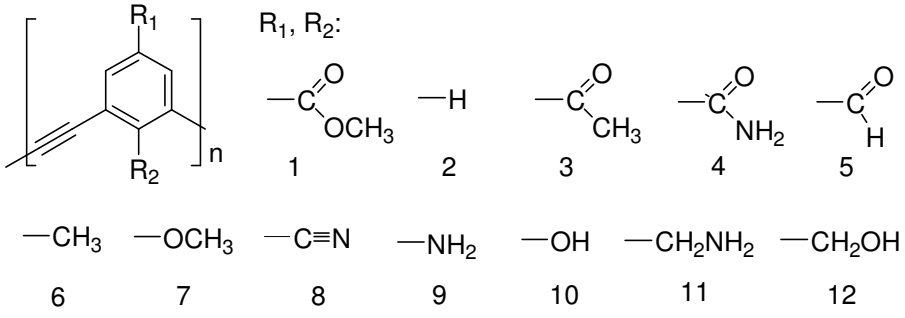


Figure 5.1 Structure of functionalized mPPEs. R_1 is an *exohelix* functional group, which would be on the outer wall of the helical polymer structure; R_2 is an *endohelix* functional group, which would be encased inside the cavity of the helical polymer structure.

Aside from their ability to change the environment inside the mPPE helix cavity, *endohelix* functional groups may also influence helix stability by enhancing the π-stacking of overlapping aromatic rings. This effect is brought about because, in general, substituted aromatic rings have stronger π-stacking interactions in comparison to benzene.[17-18] On the other hand, they could also inhibit helical formation through steric interactions. The steric effect was shown by Arnt and Tew[4-5] with an mPPE having large ether *endohelix* functional groups, and by Adisa and Bruce[7-8] in a modeling study with similar mPPE structures. Steric interactions of course depend on the size of the functional groups, and in our previous study,[14] the folding behavior of an mPPE having small *endohelix* functional groups was found to be similar to that of an mPPE with only hydrogen *endohelix* functional groups.

In this study, we examine how the stability of the mPPE helical structure is affected by *endohelix* functional groups, including those that are capable of forming hydrogen bonds. The idea of incorporating hydrogen bonds into an mPPE structure was previously explored by Cary and Moore,[19] by placing suitable functional groups at the position that is *ortho* to the ethynylene linkages as well as on the outer wall of the mPPE helical structures. In that study, hydrogen bonds between the functional groups significantly improved the stability of helical structures, allowing them to withstand a higher ratio of chlorinated solvent in solvent mixtures. Because hydrogen bonds are far stronger than the π-stacking effect, they would presumably lead to more stable helical structures and bring the mPPE secondary conformations a step closer to those of biological macromolecules, which are primarily stabilized by a series of hydrogen bonds.[20]

In an earlier study,[14] we employed a replica exchange molecular dynamics (REMD) simulation protocol to examine the effect of *exohelix* functional groups on the folding behaviors of a wide range of functionalized mPPEs. While experimental results have shown that the factors affecting secondary structure in mPPEs are so exceedingly complex that they preclude classification by simple heuristics, we found that our simulation results proved an accurate indicator of whether a given mPPE will fold under given conditions.

The excellent agreement of our previous simulation results with numerous published experimental results as well as the results from other published modeling works indicate that molecular simulation is a useful tool for exploring the folding behavior of mPPEs, presenting a cost effective alternative to laboratory synthesis and characterization. Thus, in this study, we employed our REMD protocol as well as classical molecular dynamics (MD) to study the effect of the aforementioned *endohelix* functional groups on the folding behaviors of their respective mPPEs.

5.2 Computational method

5.2.1 Computer system

All simulations were conducted using the *Palmetto* supercomputer[21] at Clemson University. The Palmetto cluster consists of many different computer systems. The computers used in this study were Dell PE 1950 with 2x Intel Xeon E5345 Quad Core

processors at 2.33 GHz, 4MB L2 Cache, 12GB RAM and 80GB of local storage (120TB of network storage). The computers and network are linked by 10G-SW32LC-16M Myrinet linecards.

REMD simulations in this study employed from 24 to 64 CPUs, and took 32 to 72 hours to complete 10 ns of simulation time with Gromacs version 3.3.1 (1 CPU for each replica).[3, 22-27] Also, significant decreases in computation time were observed with simulations using updated release of Gromacs version 4.0.5 or higher (2 or 4 CPUS for each replica).[28]

5.2.2 Simulation details

The REMD protocol employed for this work is identical to that described in Chapter 2 and in published reports from our group[6-8, 14] and, as such, is only briefly described here. All constant volume REMD and MD simulations were conducted using Gromacs, version 3.3.1 or 4.0.5,[22-29] on the *Palmetto* supercomputer[21] at Clemson University. The starting mPPE structures were built in Materials Studio 4.4,[30] then transferred to Gromacs. Isothermal conditions for all simulations were maintained using Berendsen temperature coupling (T = 300 K, τ_T = 0.1 ps),[31] whereas Parrinello-Rahman pressure coupling[32] (P = 1 bar, τ_P = 1 ps) was used to maintain system pressure during the 200 ns isothermal-isobaric simulations that were used to equilibrate the solvated mPPE systems prior to REMD and MD simulations. The LINCS algorithm was employed to hold constant all

mPPE bond lengths.[24] Finally, Visual Molecular Dynamics was used to render graphics of mPPE structures.[33]

To evaluate the folding behaviors of a given mPPE system, we used Gromacs to conduct two REMD simulations, one initialized with the mPPE in a random extended structure and the other with the polymer in a helical conformation. Results were used to categorize the various mPPEs into groups, based on the time evolution of the radius of gyration (R_g) observed during simulations of the polymers. These groups are designated by a number, 1 to 4, which indicates the likelihood that the mPPE will form a stable helix in a given solvent (Group 1 being least likely and Group 4 being most likely). Refer to Chapter 2 for a detailed description of the method and classification system. So as to monitor the mPPE conformation changes during the simulation, we employed several different parameters including the radius of gyration, the Lennard-Jones interactions between mPPE and solvent molecules, the Lennard-Jones interactions between mPPE atoms, the solvent accessible surface area, the average distances between overlapping aromatics rings as well as visual inspection of the trajectories. All of these parameters showed similar correlation with the folding process. Thus, only the time evolution of the radius of gyration is reported in the text.

The mPPE and solvent models were taken from the Optimized Potentials for Liquid Simulations[34-40] (OPLS), following the procedure described earlier.[6-8] The *endohelix* functional groups, numbered 6 to 10 in Figure 5.1, were selected to represent a range of size, polarity, and their ability to form hydrogen bonds (containing a hydrogen bond

donor such as Nitrogen or Oxygen, and a hydrogen bond acceptor). Further, we selected only groups whose sizes were small enough to fit inside the helix cavity, thus, avoiding any steric effects, as such interactions could severely hinder the formation of the helical structure.[4-5, 7-8] Three explicit solvents were used for this study: acetonitrile, methanol, and chloroform. Hydrogen bond analyses were performed using Gromacs, with a cut off distance of 0.35 nm and an angle cut off of 30°, as in a previous study.[41]

For this study, twelve different *exohelix* functional groups were investigated. Among them, R_1 = $-COOCH_3$ and H were fully studied for all seven *endohelix* functional groups, R_2 from 6 to 12 (see Figure 5.1). The others were only studied with R_2 = $-CH_2NH_2$ and $-CH_2OH$. REMD and MD simulations for mPPEs with R_1 = $-COOCH_3$ and H and R_2 = $-CH_2NH_2$ and $-CH_2OH$ were conducted in explicit acetonitrile (CH_3CN), methanol (CH_3OH) and chloroform (CH_3Cl). For the other mPPEs, only REMD simulations in explicit acetonitrile were conducted.

5.3 Results and discussion

In addition to the predictive value of the REMD simulations, the results of this work provide insight into the folding process of mPPEs, and of macromolecules in general. Particularly, we offer new observations for mPPEs that are capable of forming hydrogen bonds, elucidating the role that such hydrogen bonds play in the folding process. In what

follows, we present simulation results which will be of use to researchers designing mPPE architectures with tailored properties and stable helical secondary structures.

5.3.1 Secondary structure of mPPEs with ester *exohelix* functional groups ($R_1 = 1$)

The ester, *exohelix* functionalized mPPE ($R_1 = 1$) is one of the few mPPEs that maintains a stable helical structure independent of the *endohelix* functional groups attached to the aromatic rings of the polymer. This allows for control over the environment within the helical cavity without disrupting polymer structure. In this section, we report REMD simulation results for several ester-functionalized mPPE structures having different *endohelix* functional groups, including those reported in the literature.[15-16]

Simulations were conducted in explicit acetonitrile for seven mPPE systems with ester *exohelix* functional groups ($R_1 = 1$). The time evolution of R_g is shown in Figure 5.2 for each of these simulations. Among them, six mPPEs were observed to have similar folding behaviors to that of the hydrogen functionalized ester mPPE ($R_1 = 1$, $R_2 = 2$). That is, the simulations indicated that the polymers would fold into helical structures in acetonitrile. The only exception to this is the mPPE having $R_2 = -CN$, which was observed to only partially fold. This exception is similar to results from our previous simulation study[14] in which an mPPE with the same functional group in the external position ($R_1 = -CN$) only partially folded in acetonitrile and other solvents.

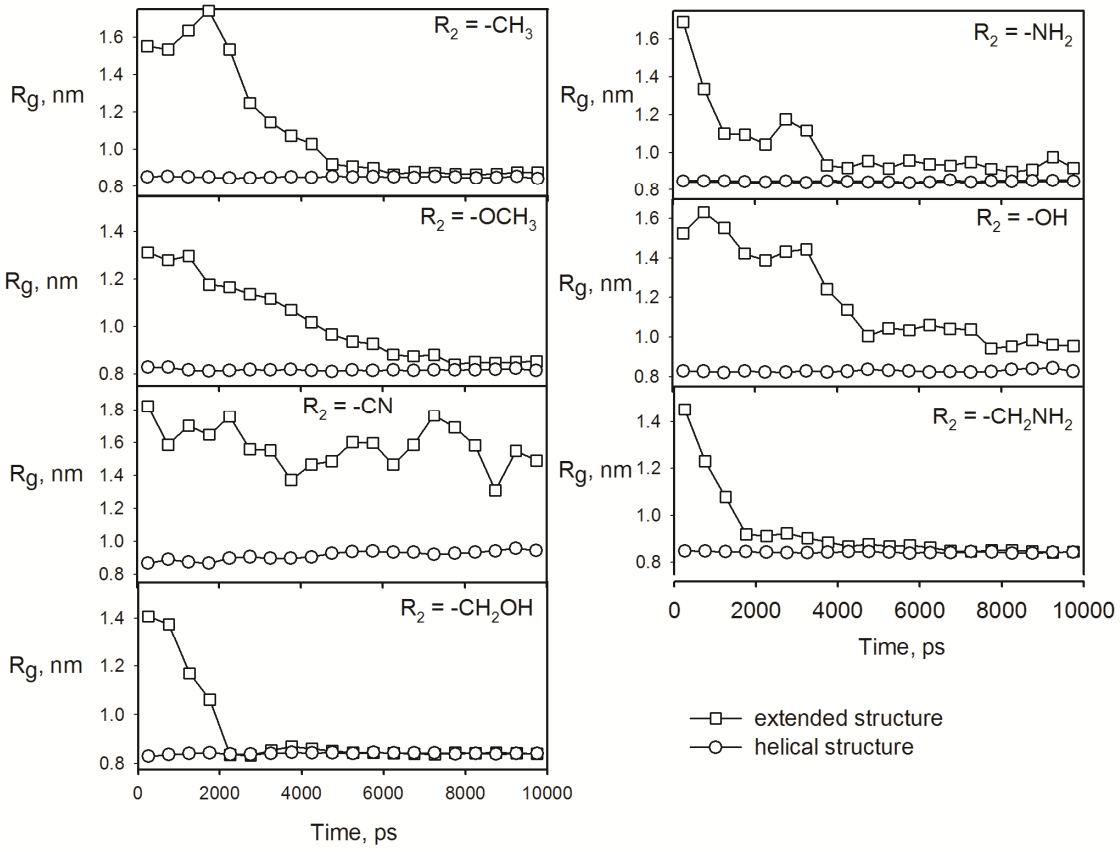


Figure 5.2 Time evolutions of the radius of gyration (R_g) for seven mPPEs in acetonitrile that contained ester *exohelix* functional groups (R_1 = –$COOCH_3$) and differing *endohelix* functional groups (R_2). Each plotted point is a time average over 500 ps, placed at the center of the corresponding time interval.

Of the *endohelix* functional groups considered in Figure 5.2, two are non-polar (R_2 = 6 and 7), one has a moderate dipole moment (R_2 = 8), and four are strongly polar (R_2 = 9, 10, 11, 12). Because each of these substituents have essentially the same results in Figure 5.2, it is reasonable to conclude that the polarities of *endohelix* functional groups are

unlikely to have a strong effect on the folding behavior of mPPEs functionalized with ester *exohelix* groups. These results further indicate that the helix cavity of an mPPE can be tailored to be polar or non-polar as needed for a given application without compromising the mPPE's ability to fold, so long as the *exohelix* functional group is an ester.

5.3.2 Hydrogen bonding in mPPEs with $R_1 = -COOCH_3$

Further analysis of the REMD simulations for mPPEs with ester *exohelix* functional groups ($R_1 = -COOCH_3$) and polar *endohelix* functional groups ($R_2 = 9$, 10, 11 or 12) showed that hydrogen bonds were formed between the *endohelix* functional groups. Specifically, two kinds of hydrogen bonds were observed: i) those between functional groups of adjacent residues, i.e., between residues numbered *n* and *n*+1. We refer to this type of interaction as an *intra-turn* hydrogen bond (intra-turn HB), an example of which is pictured in Figure 5.3a; and ii) those between functional groups of overlapping residues, i.e., between residues numbered *n* and *n*+*x*, with $x \geq 5$. Such an interaction is referred to as an *inter-turn* hydrogen bond (inter-turn HB), an example of which is depicted in Figure 5.3b.

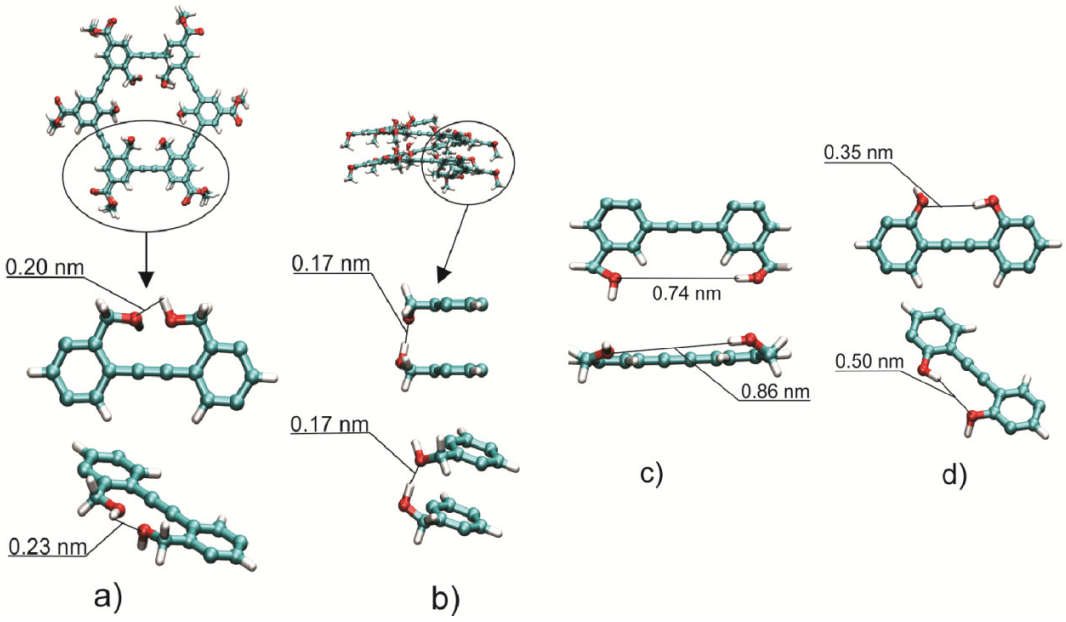


Figure 5.3 Representative distances between hydrogen bond (HB) donors and acceptors attached to the functional groups of helical mPPEs: HB distances between a) adjacent and b) overlapping residues for an mPPE with *endohelix* functional groups $R_2 = -CH_2OH$. Also pictured are HB distances for c) an mPPE with *exohelix* functional groups $R_1 = -CH_2OH$, and d) an mPPE with *endohelix* functional groups $R_2 = -OH$. Atom colors for carbon, hydrogen and oxygen are teal, white and red, respectively.

The geometries of mPPEs are such that both intra-turn and inter-turn hydrogen bonds are possible, depending on the functional groups and their positions on the phenylene rings. For example, in a typical mPPE helical conformation with *endohelix* functional groups $R_2 = -CH_2NH_2$ or $-CH_2OH$, the distances between hydrogen donors ($-NH_2$ or $-OH$) and hydrogen acceptors ($-N$ or $-O$) are approximately 0.2 to 0.25 nm for groups on adjacent phenylene rings and approximately 0.17 nm for groups on overlapping rings. These

values are within the normal range for hydrogen bonds. On the other hand, when these groups (R_1 = $-CH_2NH_2$ or $-CH_2OH$) are at *exohelix* positions, we see a donor-acceptor separation of more than 0.3 nm between adjacent rings, indicating that intra-turn HBs are not likely to form in such mPPEs (Figure 5.3c). A similar conclusion could be made for mPPEs containing *endohelix* groups R_2 = $-NH_2$ or $-OH$ (see Figure 5.3d); the distance between neighboring functional groups is too far to allow intra-turn HBs to form.

When analyzing hydrogen bond formation in REMD simulations of mPPEs containing *endohelix* functional groups R_2 = $-NH_2$ or $-OH$, only inter-turn HBs were observed, while for mPPEs having R_2 = $-CH_2NH_2$ or $-CH_2OH$, both inter-turn HBs and intra-turn HBs were present. These observations are consistent with the donor-acceptor distance analysis described above.

5.3.3 Effect of hydrogen bonds on folding behavior

Hydrogen bonds are far stronger than π-stacking interactions (van der Waals interactions), which are the primary contributor to stabilization of the helical conformation in most mPPEs. A typical non-ionic hydrogen bond in an organic molecule has a bond dissociation energy of approximately 4-5 kcal/mol,[41] which is stronger than π-stacking interactions, which commonly have a dissociation energy of from 1.5 to 2.5 kcal/mol.[18, 42-44] The strong stabilizing effect of hydrogen bonds on mPPE helical structures was experimentally confirmed by Cary and Moore for an mPPE having

hydrogen bonds formed between *exohelix* functional groups.[19] Thus, it is logical to conclude that the presence of hydrogen bonds would make the helical structures significantly more stable than those without hydrogen bonds. Further, the quick convergence of the folding process observed in simulations initialized with extended structures (e.g., shown in Figure 5.2 for $R_2 = -CH_2OH$ or $-CH_2NH_2$) suggests that hydrogen bond formation may significantly increase the rate of mPPE folding, although the Monte Carlo exchange moves incorporated into REMD prevent a quantitative comparison of kinetics from these simulations.

When examining the role that hydrogen bonds play in the folding process, several scenarios are possible. In one scenario, the hydrogen bonds form prior to the onset of folding, possibly playing a role in initializing the folding process. Alternatively, the hydrogen bonds may form consecutively as folding progresses, indicating that hydrogen bond formation is a possible driver of the folding process. A third possibility is that the mPPE helical structures form due to the usual π-stacking interactions, as they would in non–hydrogen bonded mPPEs, with the hydrogen bonds formed only after the required proximity is obtained through π-stacking. This information is easily gathered from the REMD simulations, and may be different for intra-turn HBs and inter-turn HBs.

Results from the simulations initialized with the respective mPPE in an extended structure showed that inter-turn HBs formed only while the mPPEs folded, inversely correlating with the decrease in R_g of the mPPE (as shown in Figure 5.4). This is expected, since inter-turn HBs can only form once the position advantage of an overlap is

available to them. On the other hand, for mPPEs having $R_2 = -CH_2NH_2$ or $-CH_2OH$, where both inter-turn HB and intra-turn HB hydrogen bonds were observed, results indicate the presence of intra-turn HBs prior to the onset of folding, with the number increasing as folding progresses.

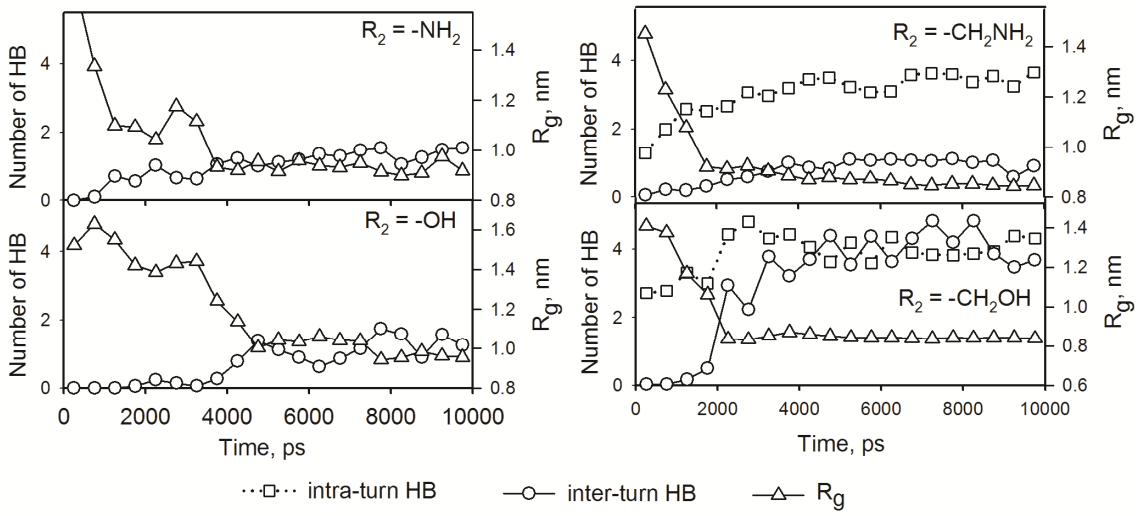


Figure 5.4 Time evolutions of the numbers of hydrogen bonds (inter-turn HBs and intra-turn HBs) during simulations initialized with extended structures for mPPEs having *exohelix* functional group $R_1 = -COOCH_3$ in acetonitrile. Plotted points are averages over 500 ps, placed at the middle of the time range. The absence of intra-turn HBs in a) and b) is expected, as the intra-turn distances between *endohelix* functional groups $R_2 = -NH_2$ or $-OH$ prohibit their formation.

By their nature, intra-turn hydrogen bonds between *endohelix* functional groups promote *cisoid* conformations within mPPE chains, which is a necessity for helix formation.

Furthermore, *cisoid* conformations are stabilized by intra-turn HBs in a single step via bond rotation about a single ethynylene linkage. This situation is considerably different in mPPEs that cannot form intra-turn HBs, because stabilization occurs only when an overlap interaction is formed, either through π-stacking of aromatic rings or through the formation of inter-turn hydrogen bonds. However, a set of coordinated rearrangements is required before the first overlap forms. Therefore, intra-turn HBs accelerate the folding process in two ways: by facilitating the initial overlap interaction, and by providing an additional driving force for the formation of each subsequent *cisoid* conformation. If mPPE folding is viewed as a nucleation/growth process, intra-turn HB interactions facilitate both nucleation events and growth, while inter-turn HBs and π-stacking interactions affect only the growth process. Therefore, the additional effect of nucleation due to intra-turn HBs is likely responsible for the enhanced folding rates evident in Figure 5.2 for $R_2 = -CH_2OH$ or $R_2 = -CH_2NH_2$.

5.3.4 Secondary structure of mPPEs without *exohelix* functional groups ($R_1 = -H$)

In order to clarify the effect of the *endohelix* functional groups on the folding behavior of mPPEs, and to explore the role of hydrogen bonds in helix formation, we conducted a series of REMD simulations for several mPPEs having only hydrogen in the *exohelix* functional group position ($R_1 = -H$). The absence of *exohelix* functional groups in these oligomers allows for a better examination of the effect of the *endohelix* functional groups on their respective mPPE folding behaviors. Further, these may be compared to previous

work[14] in which we considered mPPEs with similar functional groups in the R_1 position instead of the R_2 position. The simulation results for these mPPEs are shown in Figure 5.5 and summarized in Table 5.1.

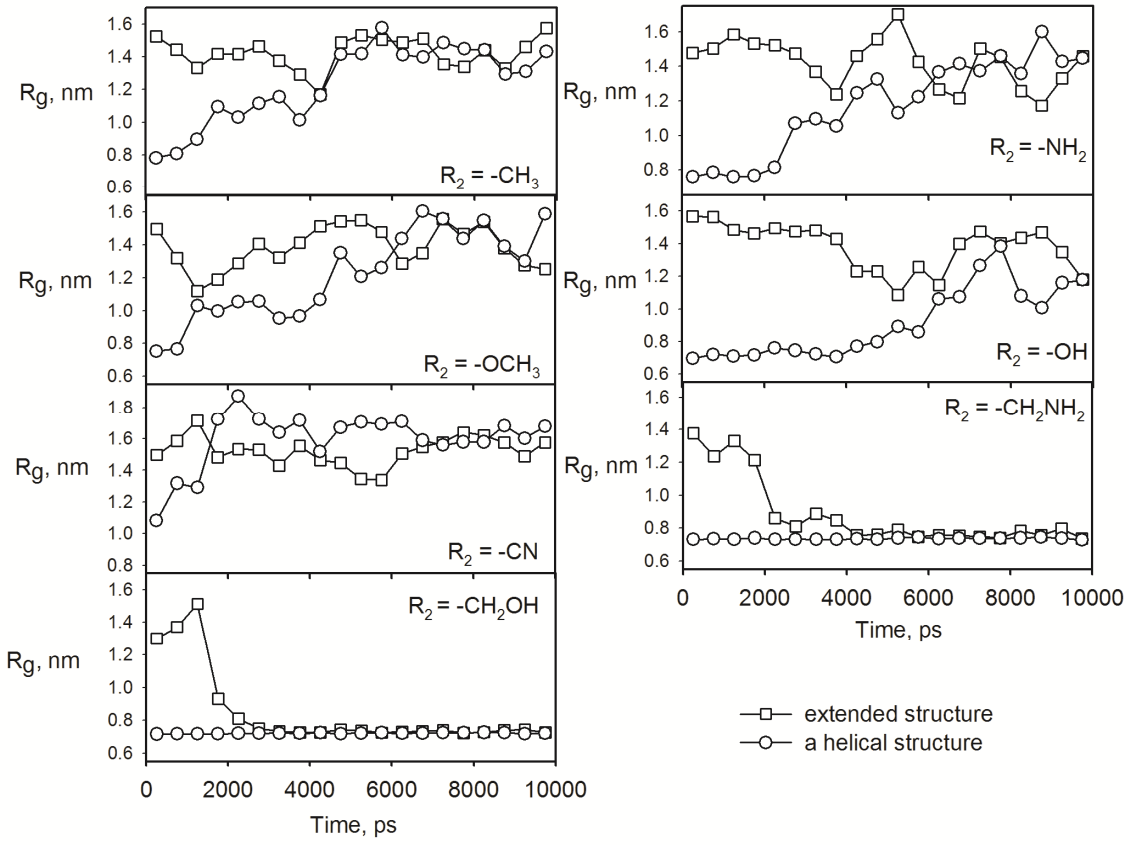


Figure 5.5 Time evolutions of the radius of gyration (R_g) for seven mPPEs having the *exohelix* functional group $R_1 =$ −H in acetonitrile but with different *endohelix* functional groups (R_2). Each plotted point is a time average over 500 ps, placed at the center of the corresponding time interval.

Simulation results showed that five of seven mPPEs, those having functional groups R_2 = 6, 7, 8, 9, and 10 at the *endohelix* positions, exhibited similar folding behaviors to those mPPEs with the same groups placed at *exohelix* positions (R_1). This indicates that the positions of the functional groups do not have a significant effect on their respective mPPE folding behaviors, suggesting that the effect of functional groups on π-stacking has stronger implications for secondary structure formation than does the placement of solvophobic and solvophilic interactions along the chains.

The results for mPPEs having $R_2 = -NH_2$ and $-OH$ demonstrate that the inter-turn HBs present in the helical mPPE structures at the beginning of the simulations were not sufficient to maintain the stability of the helices, causing the polymers to transition to less ordered conformations. Further, the absence of any helical secondary structure in simulations initialized with extended conformations demonstrates that the inter-turn HBs in these mPPEs are not able to significantly stabilize the π-stacking pairs, should any overlap occur.

As shown in Table 5.1, the two mPPEs capable of forming *endohelix* intra-turn HBs (those with $R_2 = -CH_2NH_2$ or $-CH_2OH$) have completely different folding behaviors when compared to mPPEs with these groups at the *exohelix* position R_1, where group spacing prohibits the formation of intra-turn HBs. At the *exohelix* position, these groups do not stabilize the helical conformation (i.e., the polymer does not fold into helical structures – Group 1 behavior), while at the *endohelix* position, they lead to increases in the number of *cisoid* conformations and lead to the mPPE folding into stable helical

structure (Group 4 behavior). These results again suggest that the ability to form intra-turn HBs is a significant advantage in terms of secondary structure formation, when compared to inter-turn HBs.

Table 5.1 REMD simulation results for various functionalized mPPEs in acetonitrile, where the extent of folding into helical structures is categorized into four groups: 1 (non-folding); 2 and 3 (partial folding); and 4 (fully folding).[14]

Functional group		mPPEfolding behaviors in acetonitrile[a]	
		(*endohelix*)[b] $R_1 = -H$, R_2 = specified group	(*exohelix*)[c] R_1 = specified group $R_2 = -H$
6	$(-CH_3)$	1	1
7	$(-OCH_3)$	1	1
8	$(-CN)$	1	2
9	$(-NH_2)$	1	1
10	$(-OH)$	2	1
11	$(-CH_2NH_2)$	4	1
12	$(-CH_2OH)$	4	1

[a] The reported mPPE structures having specified functional groups either at *endohelix* (R_2) or *exohelix* (R_1) positions. Hydrogens were placed at the remaining functional group positions.

[b] Based on REMD simulation results in this study.

[c] Based on results in a previous study[14] (presented in Chapter 2).

For further investigation of the effect of intra-turn HBs on secondary structure formation, REMD simulations were conducted for mPPEs having $R_2 = -CH_2NH_2$ or $-CH_2OH$ in two additional solvents. We chose to examine their behavior in methanol, as this solvent

is capable of forming hydrogen bonds that may disrupt the intra-turn HBs, and in chloroform, which is known to deter helical structure formation with most mPPEs. The results are given in Figure 5.6, showing that the mPPEs were indeed able to fold in both methanol and chloroform. Although the transitions from extended conformations to helical conformations in chloroform were slower than those observed in acetonitrile or in methanol, results clearly show that the helical structure is stable and that the extended structure has a tendency to fold into the helix. Should this prediction be correct, we feel that an mPPE with the ability to fold in chloroform is a significant finding, since chloroform generally acts as a denaturing solvent for most known mPPEs.[2, 14, 45] Further, the polymers often have low solubility in solvents that promote folding into helical structures, and this result shows that it may be possible to obtain an mPPE/solvent system that exhibits high mPPE solubility in addition to promoting mPPE secondary structure (or helical structure) formation. Our result is not without precedent. An earlier study by Cary and Moore showed that mPPE helices stabilized by hydrogen bonds are much more resistant to denaturation than similar mPPEs without hydrogen bonds.[19] Our finding suggests that additional hydrogen bonds could strongly stabilize the mPPE helical structure even in denaturing solvents, such as chloroform.

5.3.5 Analysis of hydrogen bonds for R_2 = 11 and 12

Time averages of the number of hydrogen bonds during REMD simulations are shown in Table 5.2. The average number of hydrogen bonds in mPPEs with $R_2 = -CH_2OH$ are

always higher than those in mPPEs with $R_2 = -CH_2NH_2$, consistent with the higher level of electronegativity of oxygen (3.44) as compared to nitrogen (3.04).[46] The electronegativity values also reflect the strength of the respective hydrogen bonds; specifically, hydrogen bonds between -OH and O are stronger than those between −NH and N. In most cases, we observe that the average numbers of intra-turn HBs are higher than those for inter-turn HBs. This latter observation results from the inter-turn HBs being dependent on the distances between overlapping aromatic rings, meaning any disturbance in these distances could easily lead to inter-turn HB breaking. There is one exception to this rule, observed for mPPEs having $R_2 = -CH_2OH$ in acetonitrile. For this case, there are more inter-turn HBs than intra-turn HBs, indicating the hydrogen bonds of mPPE helical structures are less likely to be disturbed than in methanol or chloroform.

When the *exohelix* (R_1) functional groups attached to these polymers are changed from –H to $-COOCH_3$, there is a significant change in the number of hydrogen bonds formed between the *endohelix* (R_2) functional groups in acetonitrile, methanol, and chloroform. Additionally, in acetonitrile, which does not have the ability to form hydrogen bonds, and chloroform, which only has hydrogen acceptor, the difference between the average number of intra-turn HB and inter-turn HB is less apparent. On the other hand, in methanol, we observed a clear decrease in the number of inter-turn HBs, likely resulting from the competition between intramolecular hydrogen bonds between mPPE functional groups and intermolecular hydrogen bonds between mPPE functional groups and solvent molecules. When the *exohelix* functional group is set to hydrogen (R_1 = -H), this same

trend is observed for the mPPE having R_2 = -CH_2OH. However, the mPPE having R_1 = -H and R_2 = -CH_2NH_2 exhibited the opposite behavior, with the number of mPPE hydrogen bonds in methanol being slightly greater than in the other solvents.

Table 5.2 Hydrogen bond (HB) analysis from MD simulations of four mPPEs (R_1 = –$COOCH_3$ or –H, R_2 = –CH_2NH_2 or –CH_2OH) in three solvents: acetonitrile (CH_3CN), methanol (CH_3OH), and chloroform (CH_3Cl).

Property	mPPE, R_1 = -$COOCH_3$, R_2 = -CH_2NH_2			mPPE, R_1 = -$COOCH_3$, R_2 = -CH_2OH		
	Solvent			Solvent		
	CH_3CN	CH_3OH	CH_3Cl	CH_3CN	CH_3OH	CH_3Cl
Total HB[a]	4.5 ± 1.7	4.0 ± 1.6	4.5 ± 1.8	8.3 ± 1.5	4.0 ± 1.6	8.0 ± 1.8
intra-turn HB	3.4 ± 1.4	3.2 ± 1.3	3.5 ± 1.4	4.0 ± 1.2	3.3 ± 1.4	4.8 ± 1.5
inter-turn HB	1.0 ± 1.0	0.7 ± 0.8	1.0 ± 1.0	4.1 ± 1.9	0.5 ± 0.8	2.8 ± 2.1
HB bond length (nm)[b]	0.22	0.22	0.22	0.20	0.20	0.20
HB average angles (°)[c]	16.1 ± 7.4	16.0 ± 7.4	15.5 ± 7.4	13.1 ± 7.0	13.0 ± 6.9	13.0 ± 6.8
Fold-time (ns)[d]	10	8	<5	17	35	10
Property	mPPE, R_1 = -H, R_2 = -CH_2NH_2			mPPE, R_1 = -H, R_2 = -CH_2OH		
	Solvent			Solvent		
	CH_3CN	CH_3OH	CH_3Cl	CH_3CN	CH_3OH	CH_3Cl
Total HB[a]	2.5 ± 1.3	3.1 ± 1.4	2.8 ± 1.5	7.5 ± 1.7	5.0 ± 1.8	6.3 ± 1.9
intra-turn HB	1.5 ± 1.1	2.8 ± 1.3	2.5 ± 1.3	3.1 ± 1.0	3.2 ± 1.2	4.6 ± 1.4
inter-turn HB	1.0 ± 0.9	0.2 ± 0.5	0.2 ± 0.5	4.4 ± 1.6	1.7 ± 1.4	1.6 ± 1.7
HB bond length (nm)[b]	0.23	0.22	0.23	0.20	0.20	0.21
HB average angles (°)[c]	18.3 ± 7.4	15.7 ± 7.4	15.2 ± 7.4	13.9 ± 7.0	13.2 ± 6.9	13.2 ± 6.8
Fold-time (ns)[d]	23	12	10	18	9	<5

[a] Average of total number of hydrogen bonds, including all types of intra-turn HBs and inter-turn HBs, but does not account for mPPE-solvent hydrogen bonding. Data are mean ± standard deviation.
[b] Standard deviation 0.02.
[c] HB angles are reported as donor–hydrogen–acceptor.
[d] Time at which the first helical structure was observed during MD simulations.

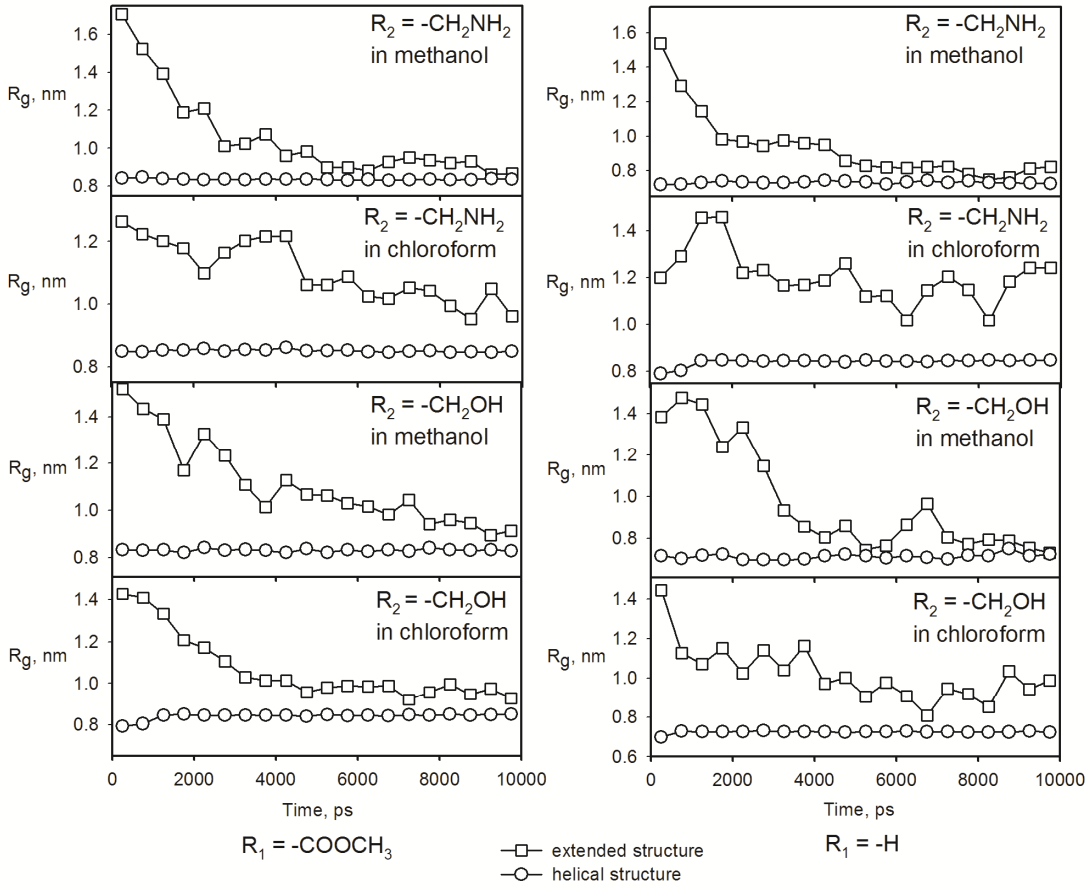


Figure 5.6 Time evolutions of radius of gyration (R_g) for mPPEs having *endohelix* functional groups $R_2 = -CH_2NH_2$ or $-CH_2OH$ in methanol and in chloroform. Plots on the left are for mPPEs with *exohelix* functional groups $R_1 = -COOCH_3$, while plots on the right are for mPPEs with *exohelix* functional groups $R_1 = -H$. Each plotted point is a time average over 500 ps, placed at the center of the corresponding time interval.

5.3.6 Effect of hydrogen bonds on mPPE folding kinetics

To determine the effect of hydrogen bonds on folding kinetics, which cannot be obtained from REMD simulations, we conducted MD simulations initialized with extended polymer structures (i.e., random mPPE conformations). These were carried out in three explicit solvents: acetonitrile, chloroform and methanol. During these MD simulations, helical structures were observed after significantly shorter times than those reported in earlier studies (see Table 5.2).[10-11] In all cases studied, the mPPEs folded into a helical conformation in less than 35 ns, even less than 5 ns in several cases, especially for simulations in chloroform. Though these simulations were not repeated for multiple initial conformations of the uncoiled polymers, we do believe that the observed folding times are representative of average values based upon observations from the REMD simulations. Further, the helical structures formed in each simulation were stable for the remainder of the simulation without any unfolding event, even when the total time was extended up to 100 ns, demonstrating that the helical structure is highly stable for the hydrogen bonding mPPEs in this study.

5.3.7 Effect of hydrogen bonds on the geometry of mPPE helical structures

Comparing the helical structures of mPPEs that form intramolecular hydrogen bonds to those that do not, there is an apparent difference: the former show much less regularity in the overlap of aromatic rings on adjacent turns as compared to the latter (see Figure 5.7).

The regular helical structures formed by mPPE systems lacking the ability to form intramolecular hydrogen bonds are stabilized by a series of π-stacking interactions, leading to conformations in which the aromatics rings overlap directly upon one another. On the other hand, the mPPE systems that readily form intramolecular hydrogen bonds often do not maintain perfect overlap of the aromatic rings on adjacent turns. This results from it being more energetically favored for the mPPE to form hydrogen bonds between neighboring functional groups than to orient the polymer backbone in such a way as to optimize the weak π-stacking interactions between overlapping aromatic rings. In these polymers, a slight structural variation in the helical arrangement of the polymer backbone can significantly increase the number of intramolecular hydrogen bonds, which leads to a lower energy mPPE conformation and simultaneously maximizes the stability of the helical structure.

The difference in the mPPE helical structures with and without intramolecular hydrogen bonds is apparent when comparing the geometries of the two polymer types. The distance between adjacent helical turns is approximately 0.37 nm for the hydrogen bonded mPPE polymers, whereas the inter-turn distance increases to nearly 0.45 nm for mPPEs unable to form intramolecular hydrogen bonds.[14] While the non–hydrogen bonded mPPEs have exactly six aromatics rings per turn, the studied hydrogen bonded structures have slightly less than this. Such a change in pitch allows for the formation of a small number of inter-turn HBs between functional groups attached to residues numbered *n* and *n*+5 along the chain, whereas a non–hydrogen bonded mPPE with six rings per turn has π-stacking interactions only between residues numbered *n* and *n*+6. The internal diameter of the

hydrogen bonded helical mPPEs is 1.27 nm on average, as measured by the distance between two *endohelix* carbon atoms of two opposite aromatics rings in the helix. This is smaller than the internal diameter of the helical structures without hydrogen bonds, 1.35 nm, giving a further indication of the more compact structure for hydrogen bonded polymers.

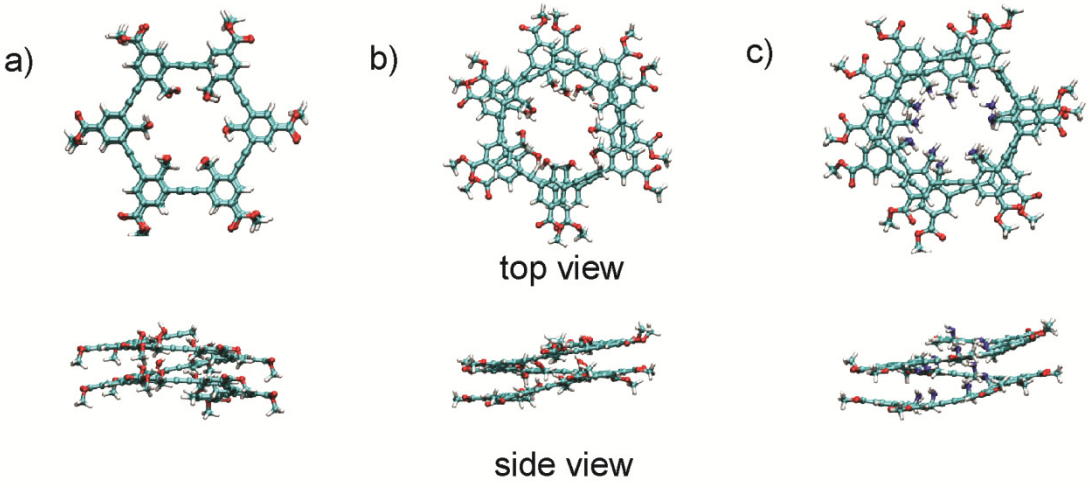


Figure 5.7 Representative helical structures of mPPEs functionalized with hydrogen bonding groups: a) a hypothetical helical structure for an mPPE with *exohelix* functional groups $R_1 = -COOCH_3$ and *endohelix* functional groups $R_2 = -CH_2OH$, where there is optimal overlap of aromatic rings from adjacent turns; b) the energy minimized helical structure from REMD for the same oligomer as in (a); and c) the energy minimized structure from REMD for an mPPE with *exohelix* functional groups $R_1 = -COOCH_3$ and *endohelix* functional groups $R_2 = -CH_2NH_2$. Atom colors for carbon, hydrogen and oxygen are teal, white and red, respectively.

5.3.8 REMD simulations for hydrogen bonded mPPEs with other *exohelix* functional groups

As shown in our REMD results for mPPEs with *endohelix* functional groups R_2 = $-CH_2NH_2$ (11) or $-CH_2OH$ (12), the presence of hydrogen bonds can significantly stabilize the helical structure of mPPEs, such that they will fold under conditions in which folding would otherwise not be expected to occur. Thus, in this section, we revisit some of our previous simulations studies,[14] where we considered different *exohelix* functional groups (R_1), but will now incorporate the *endohelix* functional groups R_2 = 11 and 12 in those polymers. These simulations were performed to ascertain whether hydrogen bonded *endohelix* functional groups can induce folding in mPPEs having *exohelix* functional groups that are known to inhibit folding.

REMD simulations of nine mPPEs (*exohelix* functional groups R_1 = 3 to 12, *endohelix* functional groups R_2 = $-CH_2NH_2$ or $-CH_2OH$ as shown in Figure 5.1) were conducted in explicit acetonitrile. Nearly identical results were observed for all of these mPPEs - all of the mPPEs were predicted to fold into stable helices (Group 4). These results suggest that hydrogen bonded mPPEs may be modified using different *exohelix* functional groups, allowing them to be compatible with a wide range of solvents, while maintaining their stable helical structure.

5.4 Conclusions

The simulation results in this work present several insights into secondary structure formation in mPPEs, which may also have significant impact on the study of secondary structure in macromolecules in general. Our results indicate that the folding behavior of a given mPPE is relatively unaltered by moving its external functional groups to positions inside the helical core. This suggests that the placement of solvophobic and solvophilic sites on the polymer is less important for mPPEs secondary formation than the effect of aromatic substituents on π-stacking interactions. This finding differs substantially from our current understanding of secondary structure in biopolymers, for example, in which solvophobic/solvophilic interactions play a major role.

Our simulations also identified two general approaches that will aid in designing mPPEs for specific applications. First, we demonstrated that choosing an *exohelix* functional group with a known propensity to induce folding will result in considerable freedom when choosing an *endohelix* functional group. When an *exohelix* the ester group was used ($R_1 = 1$), results showed that small *endohelix* functional groups (R_2) did not significantly change the folding behavior of their respective mPPEs. This indicates the feasibility of tailoring the environment inside the cavity of an mPPE, whether in terms of polarity, size, or other factors, without altering its folding behaviors. Second, we found that choosing hydrogen bonded *endohelix* functional groups (R_2) gives the same level of freedom in selecting *exohelix* functional groups (R_1), without disrupting the helical secondary structure. Thus, prudent selection of the appropriate *endohelix* functional groups would

ensure that an mPPE with desired *exohelix* functionality would exhibit a helical conformation in a wide range of solvents, as dictated by a given application.

5.5 References

1. Hill, D. J.; Mio, M. J.; Prince, R. B.; Hughes, T. S.; Moore, J. S., *Chem. Rev.* **2001,** *101* (12), 3893-4011.

2. Nelson, J. C.; Saven, J. G.; Moore, J. S.; Wolynes, P. G., *Science* **1997,** *277* (5333), 1793-1796.

3. Ray, C. R.; Moore, J. S., Supramolecular Organization of Foldable Phenylene Ethynylene Oligomers. In *Poly(Arylene Ethynylene)S: from Synthesis to Application*, **2005**; Vol. 177, pp 91-149.

4. Arnt, L.; Tew, G. N., *J. Am. Chem. Soc.* **2002,** *124* (26), 7664-7665.

5. Arnt, L.; Tew, G. N., *Macromolecules* **2004,** *37* (4), 1283-1288.

6. Adisa, B.; Bruce, D. A., *J. Phys. Chem. B* **2005,** *109* (15), 7548-7556.

7. Adisa, B.; Bruce, D. A., *J. Phys. Chem. B* **2005,** *109* (42), 19952-19959.

8. Adisa, B. Synthesis, characterization and modeling of novel polymers and polymer-templated mesoporous materials. Ph.D. Dissertation, Clemson University, Clemson, SC, U.S.A., **2005**.

9. Elmer, S. P.; Pande, V. S., *J. Phys. Chem. B* **2001,** *105* (2), 482-485.

10. Elmer, S. P.; Pande, V. S., *J. Chem. Phys.* **2004,** *121* (24), 12760-12771.

11. Elmer, S. P.; Park, S.; Pande, V. S., *J. Chem. Phys.* **2005,** *123* (11), 114902(0-14).

12. Elmer, S. P.; Park, S.; Pande, V. S., *J. Chem. Phys.* **2005,** *123* (11), 114903(0-7).

13. Lee, O. S.; Saven, J. G., *J. Phys. Chem. B* **2004,** *108* (32), 11988-11994.

14. Nguyen, H. H.; McAliley, J. H.; Batson, W. A.; Bruce, D. A., *Macromolecules* **2010,** *43* (14), 5932-5942.

15. Tanatani, A.; Hughes, T. S.; Moore, J. S., *Angew. Chem. Int. Ed.* **2002,** *42* (2), 325-328.

16. Prince, R. B.; Okada, T.; Moore, J. S., *Angew. Chem. Int. Ed.* **1999,** *38* (1-2), 233-236.

17. Sherrill, D. C., Computations of Noncovalent pi-interactions. In *Rev. Comput. Chem.*, Lipkowitz, K. B.; Boyd, D. B., Eds. Wiley: New York, N.Y., **2009**; Vol. 26, pp 1-30.

18. Sinnokrot, M. O.; Sherrill, C. D., *J. Phys. Chem. A* **2003,** *107* (41), 8377-8379.

19. Cary, J. M.; Moore, J. S., *Organic Letters* **2002,** *4* (26), 4663-4666.

20. Lehninger, A. L.; Nelson, D. L.; Cox, M. M., *Lehninger principles of biochemistry*. 5th ed.; W.H. Freeman: New York, **2008**; pp 113-148, 271-301.

21. *High performance computing*. http:/citi.clemson.edu/hpc (accessed 08/03/2011).

22. Bekker, H.; Dijkstra, E. J.; Berendsen, H. J. C., *Supercomputer* **1993,** *10* (2), 4-10.

23. Berendsen, H. J. C.; van der Spoel, D.; Vandrunen, R., *Comput. Phys. Commun.* **1995,** *91* (1-3), 43-56.

24. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J., *J. Comput. Chem.* **1997,** *18* (12), 1463-1472.

25. Hukushima, K.; Nemoto, K., *J. Phys. Soc. Jpn.* **1996,** *65* (6), 1604-1608.

26. Lindahl, E.; Hess, B.; van der Spoel, D., *J. Mol. Model.* **2001,** *7* (8), 306-317.

27. van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C., *J. Comput. Chem.* **2005,** *26* (16), 1701-1718.

28. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E., *J. Chem.Theory Comput.* **2008,** *4* (3), 435-447.

29. van der Spoel, D.; Lindahl, E.; Hess, B.; van Buuren, A. R.; Apol, E.; Meulenhoff, P. J.; Tieleman, D. P.; Sijbers, A. L. T. M.; Feenstra, K. A.; van Drunen, R.; Berendsen, H. J. C. GROMACS USER MANUAL version 4.5.4. www.gromacs.org/documentation (accessed 08/03/2011).

30. *Materials Studio*, 4.4; Accelrys Software Inc: San Diago, CA, **2008**.

31. Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R., *J. Chem. Phys.* **1984,** *81* (8), 3684-3690.

32. Parrinello, M.; Rahman, A., *J. Appl. Phys.* **1981,** *52* (12), 7182-7190.

33. Humphrey, W.; Dalke, A.; Schulten, K., *J. Mol. Graphics* **1996,** *14* (1), 33-38.

34. Jorgensen, W. L.; Madura, J. D.; Swenson, C. J., *J. Am. Chem. Soc.* **1984,** *106* (22), 6638-6646.

35. Jorgensen, W. L., *J. Phys. Chem.* **1986,** *90* (7), 1276-1284.

36. Jorgensen, W. L.; Maxwell, D. S.; TiradoRives, J., *J. Am. Chem. Soc.* **1996,** *118* (45), 11225-11236.

37. Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L., *J. Phys. Chem. B* **2001,** *105* (28), 6474-6487.

38. Price, M. L. P.; Ostrovsky, D.; Jorgensen, W. L., *J. Comput. Chem.* **2001,** *22* (13), 1340-1352.

39. Rizzo, R. C.; Jorgensen, W. L., *J. Am. Chem. Soc.* **1999,** *121* (20), 4827-4836.

40. Watkins, E. K.; Jorgensen, W. L., *J. Phys. Chem. A* **2001,** *105* (16), 4118-4125.

41. van der Spoel, D.; van Maaren, P. J.; Larsson, P.; Timneanu, N., *J. Phys. Chem. B* **2006,** *110* (9), 4393-4398.

42. Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D., *J. Am. Chem. Soc.* **2002,** *124* (36), 10887-10893.

43. Sinnokrot, M. O.; Sherrill, C. D., *J. Am. Chem. Soc.* **2004,** *126* (24), 7690-7697.

44. Tsuzuki, S.; Fujii, A., *Phys. Chem. Chem. Phys.* **2008,** *10* (19), 2584-2594.

45. Lahiri, S.; Thompson, J. L.; Moore, J. S., *J. Am. Chem. Soc.* **2000,** *122* (46), 11315-11319.

46. Lide, D. R.; Haydes, W. M. CRC handbook of chemistry and physics (Internet edition), 9-77, http://www.hbcpnetbase.com/ (accessed 08/02/2011).

CHAPTER 6

WATER DIFFUSION THROUGH *meta*-POLY(PHENYLENE ETHYNYLENE) CHANNELS

6.1 Introduction

The controlled transport of water through nanostructures is vital to the performance of many biological systems and nano-scale devices.[1-13] Therefore, the development of novel structures that can precisely control, i.e. selectively block or allow, water flow through a nano channel could lead to advances in chemical separations and sensing as well as biological applications ranging from biocides to drug delivery. One of the ways to alter water transport in a nanochannel is to coat the interior walls of the channel with functional groups, such as hydrophobic or hydrophilic species, which enhance or inhibit the wettability of the inner pore surface to water. The extent to which water can wet the surface will often dictate the relative number of water molecules in the pore, which in turn impacts the rate of water transport through the pore.[1-2, 4, 7, 9-13] However, due to the small size of many nanochannels, with diameters on the order of sub-nanometer to tens of nanometers, chemically altering the interior environment of an existing channel is often a formidable challenge. The *meta*-poly(phenylene ethynylene) (mPPE) class of materials,[14-17] on the other hand, offers an interesting platform for customizable nanochannels. Many mPPEs have the ability to self-assemble into helical structures with effective interior diameters ranging from 0.4 to 0.8 nm. Through simple variations in the monomer

chemistry, these molecules can be synthesized such that they are uniformly coated with different functional groups on both the inside and outside of the channel walls. Further, because the mPPE channel formation only takes place under specific, controllable conditions, the mPPE molecules can be easily functionalized in their denatured state, before being induced to self-assemble into nanochannels with the desired interior and exterior properties. Such a conceptual scheme is shown in Figure 6.1.

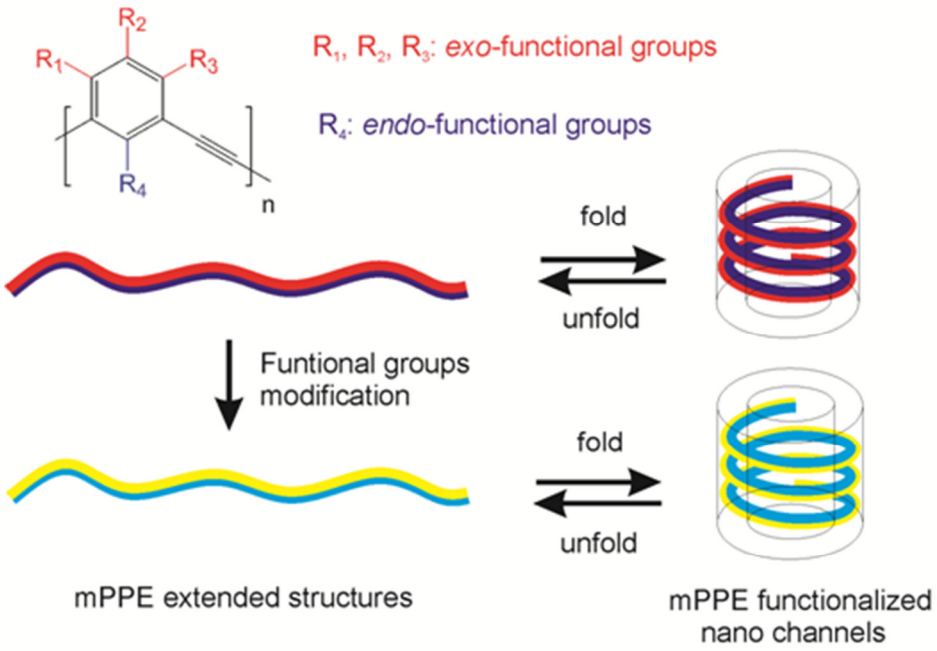


Figure 6.1 General synthetic scheme for functionalizing the interior and exterior walls of mPPE nanochannels. The colors along the mPPE chains represent the unique *exohelix* (R_1, R_2, R_3: red or yellow) and *endohelix* (R_4: purple or cyan) functional groups on the mPPE. The translucent concentric cylinders (shown on the right) define the overall shape of the mPPE nanochannels.

This concept could eliminate the challenge of conducting the functionalization process inside the confined space of an existing nanochannel, and allows for the possibility of efficiently synthesizing a wide variety of nanochannels having specifically designed selectivity - not only for water, but also for other small molecules.

There are several results from previously published experimental works[15-27] that illustrate the advantages of using mPPE helical structures as a platform for producing functionalized nanochannels. Firstly, the environment inside the cavity of mPPE helical structures (or the environment inside the channel) can be precisely tailored by selecting appropriate *endohelix* functional groups, which become localized inside the helix cavity upon self-assembly. This allows for the synthesis of porous channels having desired cross-sectional area, channel polarity and specific chemical functionality. Thus, mPPE channels have the potential to be highly selective for water, ions, or other small molecules. This opens the possibility of designing mPPE channels with a particular selectivity, for example, allowing only select chemicals to enter or diffuse through the pore in a given application. Several types of *endohelix* functional groups have been successfully incorporated into mPPE structures, including methyl,[23-24] methoxy,[17, 20] and cyano[23] functional groups. A second advantage of mPPEs for nanochannel applications is that the *exohelix* functional groups, including the solvophilic pendants, could be altered to achieve compatibility with different working environments. This latter feature is especially importance for many biological and sensor applications. A few examples of *exohelix* functionalized mPPEs included: 1) the polymers synthesized by Brunsveld et al.[18] that have non-polar alkyl pendants, which are highly compatible with a lipophilic

environment; 2) the polar ether mPPEs Nelson et al.[16] and Lahiri et al.[15] that are soluble in several moderately polar solvents; and 3) mPPEs having water-soluble, polar pendants that were prepared by Stone and Moore[26], Tan et al.[27] and Nguyen et al.[20] Several mPPE structures having multiple *exohelix* functional groups have also been synthesized. For example, an mPPE having ester and amide *exohelix* functional groups was reported by Cary and Moore[19] while others have prepared ester functionalized mPPEs that also contained cyano or amine *exohelix* functional groups.[17, 20] Another advantage of using mPPEs as nanochannels is that the overall polymer length can be precisely controlled during the polymerization process.[16, 25] Thus, the length of the mPPE nanochannel can be tailored for a specific application.

Although the concept of using mPPE polymers in their helical conformations as nanotubes has been proposed ealier,[28] no prior published study has resolved the effects of functionalizing the exterior and interior of mPPE nanochannels in regards to their performance as selective transport facilitators. Therefore, the focus of this study is to evaluate the selectivity and overall performance of functionalized mPPEs for selective diffusion applications using molecular dynamics (MD) simulation. The proposed simulation techniques have been successfully applied to the prediction of mPPE folding behaviors,[17, 21, 29-34] and the choice of simulations over direct experimental observations is easily justified given the considerable time, effort, and cost required to synthesize and test the diffusion characteristics of various mPPE nanochannels. The results provide important information about the stability of the mPPE helical conformations under conditions where species are diffusing through the nanochannels. The results also show

how the nature of the *endohelix* functional groups impacts the diffusion selectivity of the nanochannels. In summary, this computational effort greatly expands the level of understanding of diffusion phenomena occurring within the nanochannels of helical mPPEs and provides mechanistic insights which can be exploited for a range of future materials applications requiring the selective diffusion of small molecules.

6.2 Computational method

Simulations were conducted on the Palmetto supercomputer cluster[35] at Clemson University using the GROMACS simulation package, version 4.0.5.[36-40] All atomic interactions were modeled using united-atom (for octane) or all-atom (all other species) force field- parameters from the Optimized Potential for Liquid Simulation (OPLS).[41-47] Data analyses were performed using available modules within GROMACS. Visual inspection and graphic renderings were performed using Visual Molecular Dynamics.[48]

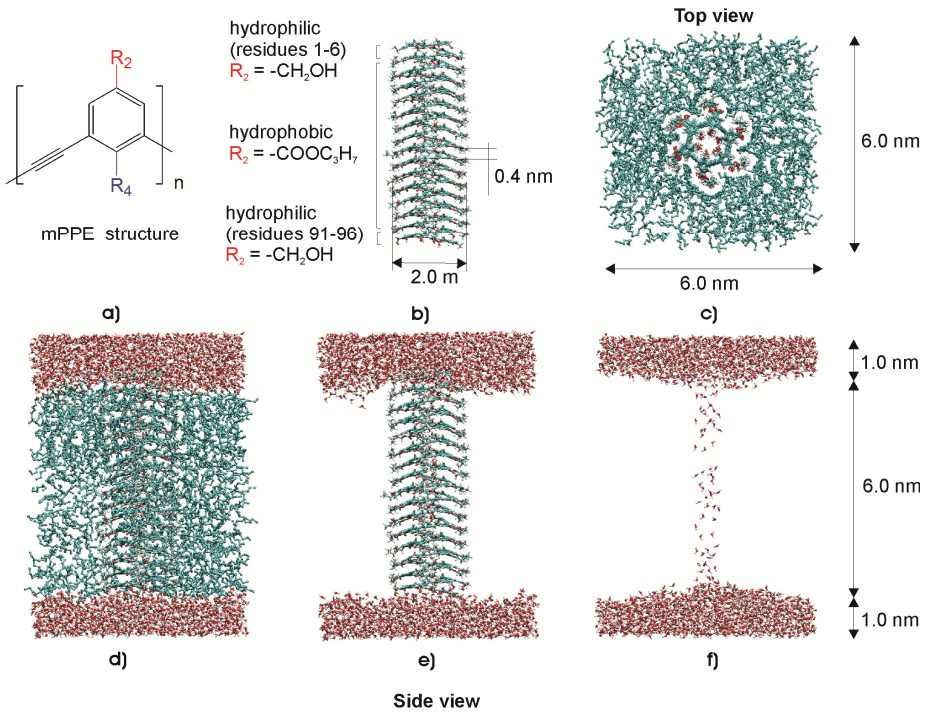


Figure 6.2 A representative mPPE helical structure and the simulation system used in this study. Depicted in the figure is the mPPE-OH nanochannel. a) a representative mPPE structure, with the *exohelix* functional groups (R_2, red color) located on the outside wall of the mPPE channels, and the *endohelix* functional groups (R_4, blue color) located on at the inside wall of the mPPE channels; b) side view of a representative mPPE channel to show the arrangement and types of *exohelix* R_2 functional groups on its outside wall; c) top view of an octane layer with an mPPE nanochannel at the center; d) a side view of a full mPPE nanochannel simulation system; e) a side view of an mPPE nanochannel simulation system without the octane layer (to clarify the position of the mPPE nanochannels); c) a side view of a simulation system without the octane layer and the mPPE nanochannel to show the distribution of the water molecules. Atom colors for carbon, hydrogen and oxygen are teal, white and red, respectively.

For our diffusion studies, we employed octane-water systems that resembled those used in simulation studies by others (see Figure 6.2).[12, 49-50] In the octane-water model, the octane layer serves as a barrier between two water layers and allows for the free movement of the mPPE nanochannels, which traverses the organic layer. The mPPE in this system mimics the function of an integral membrane protein or pore forming protein. These pore forming proteins serve many importance cellular functions and enable the flow of ions and small molecules across cellular membranes, such as phospholipid bilayers.[3, 5, 10-11, 51] Though a biological based system was chosen as the model, the diffusion characteristics within the mPPE nanochannel will generally be independent of the species surrounding the exterior wall of the helical polymer. Thus, the diffusion characteristics for the mPPEs in this study should also be valid for systems where movements of the helical mPPE are more restrictive, such as when the mPPE traverses a thin polymer film.

The simulation procedures using in this study were developed from the successful methods developed in prior mPPE simulation studies.[17, 21, 29-31] Five channels, having different interior environments, were constructed using Material Studio version 5.0[52] (see Figure 6.2 and Figure 6.3). The total length of each mPPE nanochannel was 6 nm, and the average interlayer spacing between consecutive turn of the mPPE was 0.4 nm.[17] The thickness of the water layers was set at 1.5 nm. This length was selected to represent the thickness of common lipid bilayer membranes (3 - 6 nm).[53] Ester *exohelix*-functional groups were located at the R_2 positions on the aromatic moieties (see Figure 6.2) to increase the stability of the mPPE helical structures.[15-16, 21] The ester functional groups at

R_2 monomer positions along the polymer chain had alkyl pendants, which helped to maintain the mPPE structure in the octane layer. Additionally, at both ends of the channels, the ester functional groups were replaced by polar functional groups (alkyl hydroxyl), to anchor the mPPE channel ends in the respective water layers (Figure 2). To prepare the simulation systems for water diffusion, each channel was placed in the middle of an octane-water system, with the openings in the water layers, and the body engulfed in the octane layer. Additional water molecules were used to fill the channels.

Many of the specific computational conditions used in this effort were developed in earlier studies of mPPE systems. [17, 29-31] The potential energy of each system was minimized using steepest descents followed by L-BFGS algorithms.[54] After energy minimization, each system was subjected to 2 ns of position restraint MD, in which all mPPE atoms were fixed, to allow the water and octane layers to achieve equilibrium. The final systems were used in 100 ns *semi*-NPT simulations for system stability and water diffusion analysis. The following simulation parameters were applied: v-rescale temperature coupling,[55] with a system temperature (T) of 300 K and a temperature coupling time (τ_T) of 0.1 ps; Parrinello-Rahman[56] semi-isotropic pressure coupling in the *z*-direction (main axis of the channel, as shown in Figure 6.2), with a pressure (P) of 1 atm, pressure coupling constant (τ_P) of at 1 ps, and a water isothermal compressibility of 4.5×10^{-5} bar^{-1};[57] Coulombic interactions were calculated using the smooth Particle Mesh Ewald (PME) method,[58-59] while a cut-off radius of 1.2 nm was applied for all van der Waals interactions. Finally, atom neighbor lists were updated every 10 steps and atomic trajectories were save every 10000 steps or 20 ps.

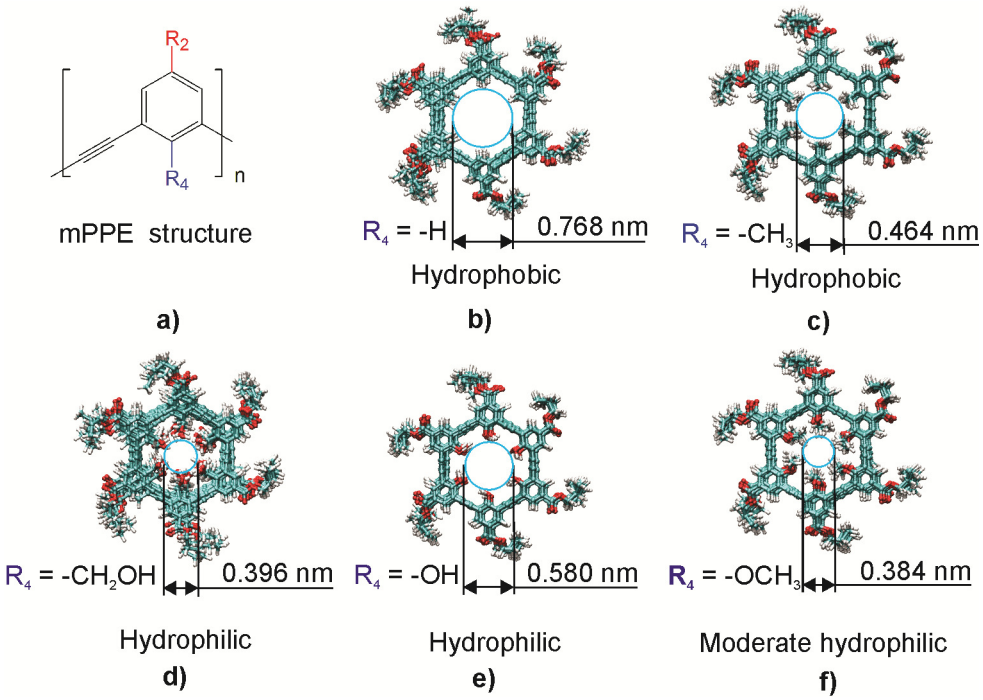


Figure 6.3 Top view of the five mPPE nanochannels examined in this study, showing the nature and polarity of the respective *endohelix* functional groups (shown as R_4 (blue) in the top left figure) as well as the effective diameter of each channel. The nanochannels shown include: a) the mPPE monomer structure; b) an mPPE-H channel; c) an mPPE-CH_3 channel; d) an mPPE-CH_2OH channel; e) an mPPE-OH channel; and f) an mPPE-OCH_3 channel. Atom colors for carbon, hydrogen and oxygen are teal, white and red, respectively.

To investigate the possibility of using *endohelix* functional groups to control the interior environment and the transport selectivity of a given polymer channel, five different mPPE channels were built (see Figure 6.3). The *endohelix* functional groups and effective diameters for each mPPE nanochannel are listed in Table 6.1. Because water transport

properties vary widely among water models, we employed five well tested water models from the literature, SPC,[60] SPC/E,[61] TIP3P,[62] TIP4P[62] and TIP5P.[63] Together with five mPPE channels, we studied a total of 25 simulation systems. The number of atoms including solvent and polymers species in each system ranged from 15000 to 20000 atoms.

The effective diameter of each channel was calculated from the simulation output by using the position of the maxima in the radial distribution function for the inner-most opposing carbon atoms of the aromatic rings, minus the van der Waals radii of the *endohelix* functional groups (see Table 6.1).[64] We note that these effective diameter calculations likely underestimate the actual values of the channel diameter, especially for those having moderately hydrophobic or hydrophilic interiors. The interaction between water molecules and the *endohelix* functional groups, especially those involving hydrogen bonding in mPPE-OH and mPPE-CH_2OH channels, could reduce the van der Waals radius of the *endohelix* functional groups, thereby increasing the effective diameter.[64] However, the effective diameters shown in Table 6.1 are sufficient for comparison purposes between mPPE nanochannels having similar interior environment, but simulated using different water models.

Table 6.1 Properties and nomenclature of simulated mPPE nanochannels.

mPPE nanochannels	*Endohelix* functional groups	Interior effective diameter (nm)[a]	Interior environment
mPPE-H	–H	0.768	Hydrophobic
mPPE-CH_3	–CH_3	0.464	Hydrophobic
mPPE-OCH_3	–OCH_3	0.384	Moderate hydrophilic
mPPE-OH	–OH	0.584	Hydrophilic
mPPE-CH_2OH	–CH_2OH	0.396	Hydrophillc

[a] Estimated using results reported by Bondi.[64]

Analysis of water diffusion behavior within channels included only those water molecules found well within the channel, clear from the opening. Water molecules at the two ends of the channel, within predefined 2.0 nm zones from the edges of the box in the *z*-direction, were excluded from the analysis to avoid anomalous end effects. Thus, the total mPPE channel length used for transport calculations was approximately 5 nm. To calculate the self-diffusivities of water inside the mPPE channels, the mean square displacements (MSD) of water molecules inside the mPPE channels in the z direction were monitored, and the water self-diffusivity inside the pore was calculated using Einstein's relation for the diffusivity. The script written for these calculations is presented in Appendix E. Specifically, the 1-D diffusivity in the *z*-direction, D_Z, was calculated using the following equation:[40]

$$\lim_{t\to\infty}(\mathrm{MSD})_z = 2D_Z t \tag{6.1}$$

where $(MSD)_z$ is the mean square displacement of water in the *z* direction and *t* is time.

The bulk three-dimensional self-diffusivity, D, of water molecules outside of the mPPE nanochannels was calculated using the following equation:40

$$\lim_{t\to\infty}(\mathrm{MSD})_{xyz} = 6\mathrm{Dt} \qquad (6.2)$$

where $(MSD)_{xyz}$ is the three dimensional mean square displacement of water molecules and *t* is time.

6.3 Results and Discussion

One of the main concerns about using mPPE helical conformations as water channels is their structural stability and integrity. Because the mPPE helical structures are formed by a self-assembly process driven by van der Waals and Coulombic forces, the resulting structures do not routinely have covalent bonds between consecutive turns of the helix. This introduces the possibility of a partial structural collapse, making the channel discontinuous between the two water layers. Additionally, the inter-turn spacing (distance between consecutive helix layers) is approximately 0.4 nm (see Figure 6.2),[16-17] thus water molecules could possibly diffuse through the channel walls or become trapped in the interlayer void spaces, further disrupting the channel structures. In the following section, the simulation data are analyzed to address these questions.

6.3.1. Stability of mPPE channels

Channels: During each 100 ns simulation, all mPPE channels remained intact, showing no indication of collapse or denaturing from the starting structures. The helical structures showed some deviation from the initial structures, as shown in the moderately winding conformation in Figure 6.4, but the variations in nanochannel geometry did not affect the continuity or integrity of the channels. The distance between overlapping aromatic rings, which could be used to evaluate the stability of the mPPE helical structures, remained stable throughout the 100 ns simulation, as shown in Figure 6.5. The average interturn distance between overlapping aromatic rings at three different time periods (0–50 ns, 25–75 ns and 50–100 ns) were almost identical, indicating a stable helical structure during the simulations. Other simulations[17, 29, 34] and experimental results[16, 20, 26, 65] have shown that similar mPPE helical structures were stable for extended period of time in both aqueous solution and nonpolar solvents (e.g. hexane).

mPPE helical structures: During the simulations, the mPPE structures were observed to drift in the *XY* plane, but their hydrophilic channel ends remained well extended into the water layers at all times. Thus, the channel design, which consists of hydrophilic *exohelix* functional groups at the terminus turns of the polymer and hydrophobic *exohelix* functional group in the middle section of the polymer, help the mPPE structure remain stable for extended time periods in the octane-water system. During a simulation with an mPPE that lacked hydrophilic *exohelix* functional groups, the mPPE structures retreated into the octane layer.

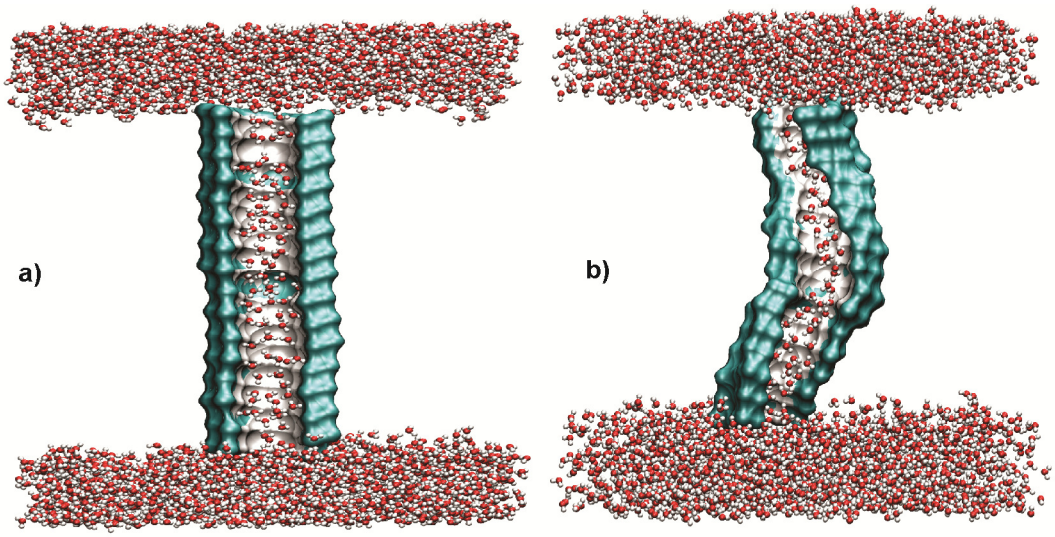


Figure 6.4 Representative surface renderings of an mPPE channel at the a) beginning and b) middle (50 ns) of a simulation. The structures shown are taken from simulation results for an mPPE-H nanochannel that employed the SPC water model. This rendering is a cutaway view, in which half of the mPPE-H atoms are omitted to reveal the interior of the channel. The octane layer, which resides between the two water layers shown, is removed for clarity. Atom colors for carbon, hydrogen and oxygen are teal, white and red, respectively.

Diffusion through the channel wall: There was initially some concern that water could diffuse through the channels walls or become trapped in the space between helical turns of the polymer. However, during all of the 100 ns simulations, no water was observed to diffuse through the wall of the channels into the octane layer.

Octane-water systems: The octane-water system also maintained a well define hydrophilic-hydrophobic interface (Figure 6.6), indicating that the simulation parameters

were reasonable and that this is an effective model system for imitating biological membrane systems, as used in previous studies.[12, 49-50] During the simulations, it was occasionally observed that a limited number of water molecules would diffuse from the aqueous layer into the octane layer. Those water molecules were monitored and excluded from the data analysis of water diffusion inside mPPE channels.

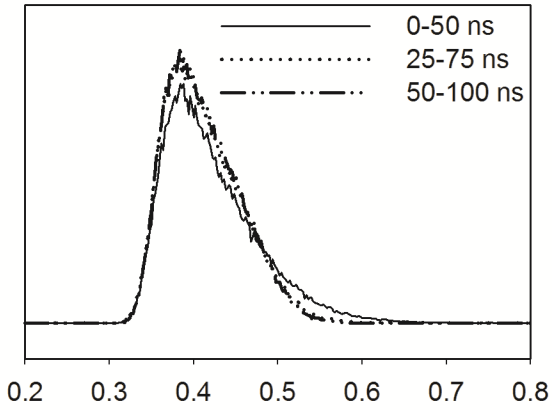


Figure 6.5 Representative radial distribution function between overlapping aromatics rings, averaged over different time periods during the simulation. The data are from the simulation of the mPPE-OH channel using the SPC water model. The significant overlap of these three lines indicates the stability of the inter-turn distance of the helical structure.

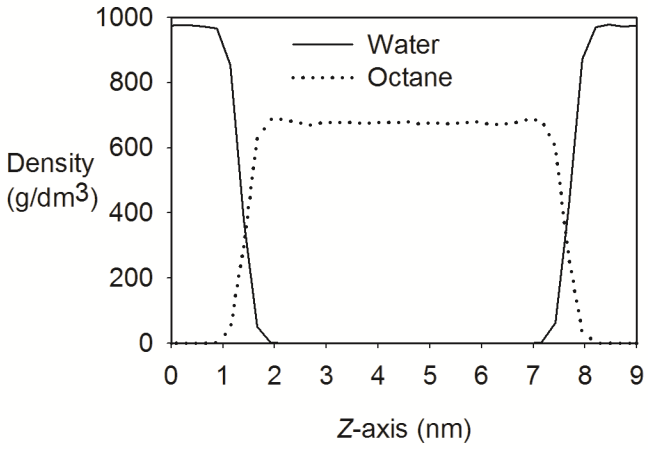


Figure 6.6 Representative octane-water system Z-axis molecular density distribution showing the octane-water interfaces that are located at approximately 1.5 and 7.5 nm. The data are from the simulation of the mPPE-OH channel using the SPC water model.

Having demonstrated that mPPE helical structures are stable conformations under the conditions of water diffusion through the self-assembled nanochannels, the following discussion addresses the issue of pore diffusion selectivity and whether water transportation can be controlled via the addition of specific *endohelix* functional groups. Specifically, we examined how the movement of water through the stable mPPE nanochannels is affected by the channel diameter and the chemical functionality of the channel interior.

6.3.2 Stability of water columns inside the mPPE channels

The five mPPE channels in this study could each be placed into one of three distinctive groups: hydrophilic channels (mPPE-OH, mPPE-CH_2OH), hydrophobic channels (mPPE-H, mPPE-CH_3, similar to hydrophobic carbon nanotubes) or moderately hydrophilic channels (mPPE-OCH_3). The effects of the different functional groups are easily shown by observing the ability of water to fill the inside of each type of channel (Table 6.2). Water, as modeled by all five water models, was able to uniformly fill, for extended periods of time, the nanochannels created by three of the functionalized helical polymers mPPE-OH, mPPE-CH_2OH and mPPE-OCH_3. Whereas, simulations with only three of the five water models predicted that water would completely fill the mPPE-H nanochannel, and no simulations, with any water model, indicated that water could accumulate or diffuse into the mPPE-CH_3 channel. Thus, the simulation results, especially those involving the mPPE-CH_3 nanochannel suggest that water diffusion selectivity of the mPPE channel can be controlled via the careful selection of *endohelix* functional groups.

The number of water molecules in the hydrophilic channels (mPPE-OH and mPPE-CH_2OH) was higher than that in the hydrophobic channel (mPPE-H), even though the estimated pore diameters of the hydrophilic channels are smaller than that of the mPPE-H channel (as shown in Table 6.1). The strong affinity of water toward the hydrophilic functional groups, as exemplified by the observed hydrogen bonds between water

molecules and the mPPE functional groups, enables water to more easily wick into the channels and increases the effective water density within the pores.

Table 6.2 The predicted average number of water molecules inside five functionalized mPPE nanochannels, using five different water models.

mPPE nanochannels	Average number of water molecules in an mPPE nano channel[a]				
	SPC	SPC/E	TIP3P	TIP4P	TIP5P
mPPE-OH	51 ± 4	58 ± 3	72 ± 6	62 ± 4	58 ± 3
mPPE-CH_2OH	30 ± 2	24 ± 3	25 ± 2	32 ± 2	33 ± 2
mPPE-OCH_3	26 ± 3	27 ± 2	26 ± 2	28 ± 2	10 ± 3
mPPE-H	45 ± 14	1 ± 2	45 ± 12	6 ± 6	37 ± 17
mPPE-CH_3	0	0	0	0	0

[a] Numbers reported are averaged over 100 ns of simulated time. Data are mean ± standard deviation.

The water diffusion simulations also revealed that the channel inside diameter is another key factor affecting the movement of water inside the mPPE channel. Interestingly, in simulations of the two hydrophilic functionalized channels, the variations in inside diameter (over the range examined) did not cause any significant change in the stability of the water columns. Whereas, simulations of the two helical polymers functionalized with hydrophobic groups showed that the ability of water to fill the channel is negatively impacted by decreases in the inside channel diameter. Specifically, three of five water models predicted a stable water column in the mPPE-H channel, whereas none of the models predicted the formation of a water column in the smaller diameter mPPE-CH_3 channel. Figure 6.7 shows the initial water distribution at the beginning of the simulation

and several snapshots of the trajectory showing the withdrawal of water out of the mPPE-CH_3 channel. Note that the water initially formed a single-file line of water, similar to those formed in carbon nanotube channel simulations.[1, 6] Further, simulations initialized with the mPPE-CH_3 channels void of any water showed that water did not enter the channel over the course of the simulation. Possible reasons for these latter observations are that the flexibility of the mPPE channel and the brush-like actions of the methyl functional groups effectively inhibit or disrupt the ability of water molecules to form ordered assemblies through the length of the channel. However, this assertion alone is insufficient to explain the expulsion of water from the polymer core because other simulations that restricted the movements of the polymer showed similar results. Thus, the inside diameter of the hydrophobic functionalized channels is in practical terms smaller than one might expect based solely on the positions of the methyl groups. Collectively, the steric and hydrophobic interactions between the water molecules and the functional groups adorning the inside surface of the mPPE channel dictate whether or not the water column is stable or unstable inside the channel.

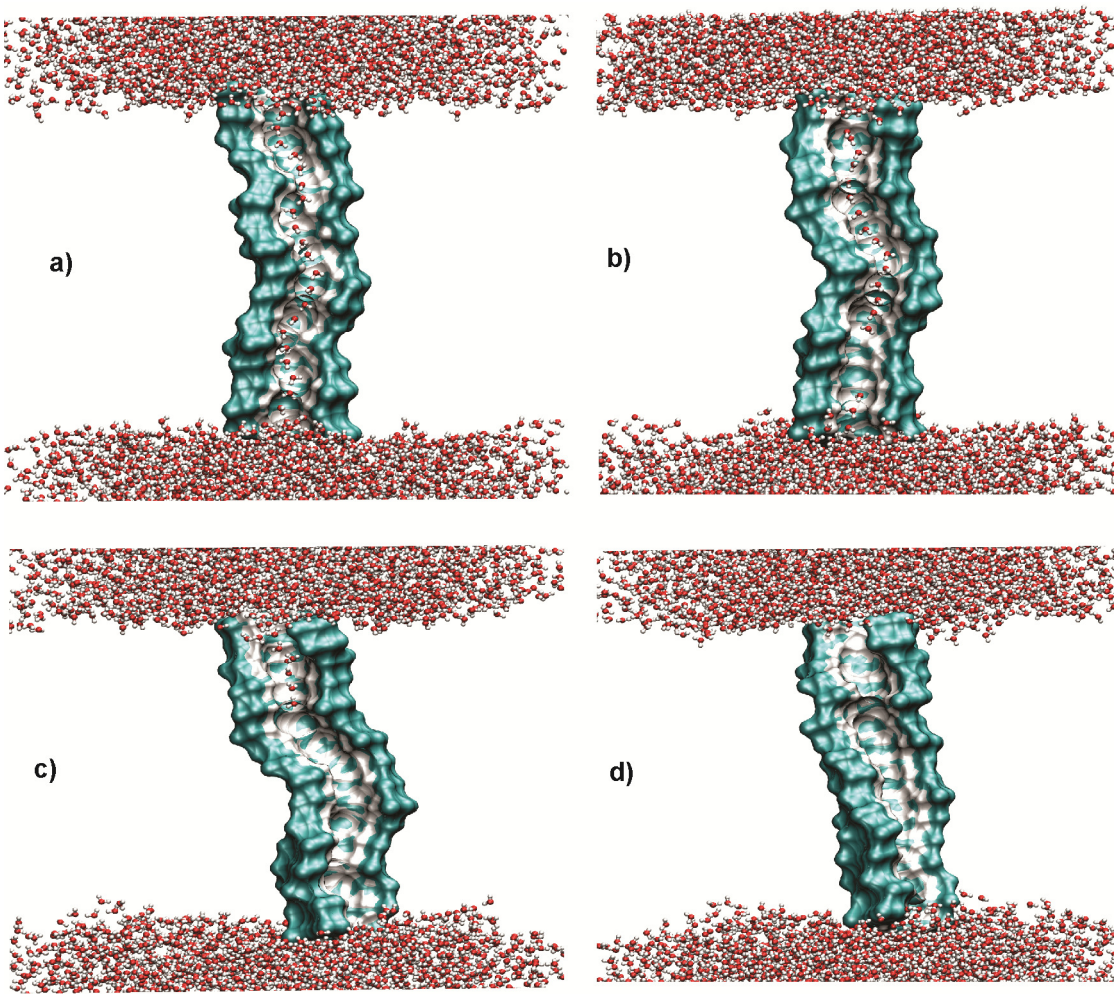


Figure 6.7 Time lapse figures showing water withdrawal from a simulated mPPE-CH_3 channel: a) water filled channel at the beginning of simulation; b) and c) the channel half-filled with water; d) a channel void of all water. The data are from a simulation of an mPPE-CH_3 channel that employed the SPC water model. Only half of the mPPE-CH_3 channel is shown so as to reveal the inside of the channel. The octane layer (between the two water layers shown) is omitted for clarity Atom colors for carbon, hydrogen and oxygen are teal, white and red, respectively.

As indicated in Table 6.2, we observed some variation in results with regard to the water models used for simulations with the mPPE-H channel. This is in contrast to simulations of the other mPPE channels listed in Table 6.1, where no differences in water diffusion behavior were observed as the water model was varied. For the water-mPPE-H system, only three (SPC, TIP3P and TIP5P) of five water models yielded simulation results that showed water could form stable water columns inside the mPPE-H channel. SPC/E and TIP4P water models transitioned from the channel being initially filled with water, to it only intermittently containing a limited number of water molecules. After the mPPE-H channel no longer contained a continuous water column, it was observed that a large number of water molecules could diffuse into the channel, but that the density of water in the channel was too low at any time to enable a contiguous water column to exist in the channel. This behavior is quite different from that observed with the smaller diameter hydrophobic channel of mPPE-CH_3, which wholly restricted the diffusion of water.

6.3.3 Water diffusion through mPPE channels

The effects of interior hydrophobicity and channel diameter of mPPE helical structures are further examined by comparing the number of water molecules diffusing into (N_T) and through (N_P) the channel, as well as the average time taken to diffuse through the channel (t_p), which is also closely correlated with the average residence time for a water molecule in the channel (see Table 6.3). The calculated water diffusivity in the *z*-

direction of the channel (D_Z) for each water model in the different mPPE channels are shown in Table 6.4.

In the hydrophobic channel mPPE-H, for simulations employing the SPC, TIP3P and TIP5P water models, a very large fraction of water molecules that entered the channel were observed to pass through the channel (9.06%, 11.08% and 6.47%, respectively), relative to the results for other mPPE channels. The average time (less than 1.5 ns) required to pass through the channel was also much lower than observed in simulations of water in hydrophilic mPPE channels. Calculation of one dimensional diffusivities showed that water molecules inside mPPE-H channels have much higher (approximately three to five times higher) diffusivities than those of water molecules inside hydrophilic channels, and their diffusivities were close to the self-diffusivity values calculated for the bulk water layers (see Table 6.4). This result is similar to the low friction water movement observed in hydrophobic carbon nanotube channels.[1, 7]

Table 6.3 The total number of water molecules diffusing into and through the simulated mPPE nanochannels for five water models.

Channels	SPC			
	N_T [a]	N_p [b]	transfer %[c]	t_p [d]
mPPE-OH	3935	14	0.35	24.3 ± 19.3
mPPE-CH_2OH	2323	50	2.15	14.4 ± 8.9
mPPE-OCH_3	1726	2	0.12	24.6 ± 34.1
mPPE-H	10186	924	9.07	1.5 ± 0.9
Channels	SPC/E			
	N_T [a]	N_p [b]	transfer %[c]	t_p [d]
mPPE-OH	3183	53	1.66	21.6 ± 12.4
mPPE-CH_2OH	1875	13	0.69	25.5 ± 20.7
mPPE-OCH_3	1247	0	0	-
mPPE-H	921	1	0.11	1.2
Channels	TIP3P			
	N_T [a]	N_p [b]	transfer %[c]	t_p [d]
mPPE-OH	7804	343	4.09	6.8 ± 4.5
mPPE-CH_2OH	2008	41	2.04	16.6 ± 9.7
mPPE-OCH_3	1082	6	0.55	37.5 ± 15.4
mPPE-H	10850	1203	11.08	1.2 ± 0.8
Channels	TIP4P			
	N_T [a]	N_p [b]	transfer %[c]	t_p [d]
mPPE-OH	3255	93	2.86	15.0 ± 10.1
mPPE-CH_2OH	2448	83	3.39	12.4 ± 7.5
mPPE-OCH_3	1342	2	0.14	37.7 ± 52.9
mPPE-H	4196	22	0.52	0.2 ± 0.2
Channels	TIP5P			
	N_T [a]	N_p [b]	transfer %[c]	t_p [d]
mPPE-OH	2118	3	0.14	0.3 ± 0.4
mPPE-CH_2OH	1567	5	0.32	58.9 ± 26.7
mPPE-OCH_3	592	5	0.84	31.2 ± 18.6
mPPE-H	7496	485	6.47	1.4 ± 0.8

[a] Total number of water molecules that diffused into an mPPE nanochannel during a 100 ns simulation.
[b] Total number of water molecules that passed completely through an mPPE nanochannel during a 100 ns simulation. Water exchange rate = $N_P \times 10^7$(molecules/s).
[c] Percentage of the number of water molecules entering the channel that passed completely through = $N_P/N_T \times 100\%$.
[d] Average transit time for a water molecule to pass completely through the channel (ns). Data are mean ± standard deviation.

In all of the simulations, the mean squared displacement (MSD) of water molecules inside the mPPE channels increased proportionally with time, allowing us to employ Einstein's relation (Eqn. 6.1) to calculate water self-diffusivities. The linear relationship between the water MSD and time in our simulations indicates a homogeneous environment for water diffusion, likely because of the uniform distribution of functional groups inside each channel and the quite stable interior channel diameters. Further, the results indicate that the time scale (10 ns) used to analyze the MSD extends well beyond the ballistic regime. Finally, the simulated diffusion behavior observed with the water-mPPE system is similar to those observed with water diffusion through carbon nano tubes that have similar interior features,[1] which further supports the observed results.

From examining the diffusion data in Table 6.3, we observe that the number of water molecules diffusing into a given mPPE nanochannel (N_T) is relatively proportional to the respective channel diameter, as listed in Table 6.1 and Figure 6.3. In contrast, the number of water molecules passing through a given channel (N_P), as well as the average transit for a water molecule to traverse the channel (t_P) is not well correlated with the inside diameter of the respective mPPE channel. In hydrophilic channels, only a small fraction of water molecules entering the channel were observed to pass completely through the channel (about 1-2%, with the largest being 5% for mPPE-OH/TIP3P), and those that did traverse the channel required a long time to do so (10–40 ns). This phenomenon and the small diffusivities calculated for these systems are largely attributable to the formation of hydrogen bonds between water molecules and the polar functional groups attached to the channel walls (see data in Table 6.4 and Figure 6.8). We also observed that water-mPPE

hydrogen bond formation most likely leads to many of the water molecules being trapped for extended periods of time at the hydrophilic (or moderately hydrophilic) sites along the mPPE channel interior (see Table 6.5). This observation is consistent with biological channel simulations by others[9] which show that water diffusion is facilitated by the interaction between water molecules and the polar functional groups adorning the biochannels; however, these interactions also lead to long water residence time in the biochannels. Though the presence of hydrogen bonding significantly slowed the diffusion of water through hydrophilic functionalized channels, it should be noted that for the mPPE-H channels, which are hydrophobic functionalized systems, no water molecule had a residence time higher than 9 ns (see Table 6.5).

Table 6.4 One dimensional (in the z-direction) diffusivity (Dz) for water inside mPPE nanochannels and bulk water self-diffusivity (D) predicted using five water models.a

mPPE nanochannels	D_z ($\times 10^{-5}$ $cm^2 \cdot s^{-1}$)[b]				
	SPC	SPC/E	TIP3P	TIP4P	TIP5P
mPPE-OH	0.35 ± 0.05	0.30 ± 0.02	0.74 ± 0.05	0.39 ± 0.05	0.17 ± 0.03
mPPE-CH_2OH	0.49 ± 0.03	0.39 ± 0.04	0.52 ± 0.02	0.50 ± 0.04	0.24 ± 0.01
mPPE-OCH_3	0.24 ± 0.04	0.17 ± 0.02	0.33 ± 0.04	0.29 ± 0.05	0.54 ± 0.10
mPPE-H	2.68 ± 0.35	-	2.78 ± 0.29	-	2.67 ± 0.20
Reported bulk diffusivity in	D ($\times 10^{-5}$ $cm^2 \cdot s^{-1}$)				
	SPC	SPC/E	TIP3P	TIP4P	TIP5P
this study[b]	4.48 ± 0.05	2.92 ± 0.04	5.77 ± 0.05	3.91 ± 0.02	2.83 ± 0.02
other studies[67-68]	4.1-5.0	2.5-3.3	5.2-7.0	3.4-4.2	2.63

[a] Experimental values for bulk diffusivity of water is 2.3×10^{-5} $cm^2 \cdot s^{-1}$ as reported by Eisenberg and Kauzmann.[69]
[b] Data are mean ± standard deviation.

Table 6.5 Highest observed residence times for water molecules inside mPPE nanochannels during simulations for five water models.

mPPE nanochannels	Residence time (ns)				
	SPC	SPC/E	TIP3P	TIP4P	TIP5P
mPPE-OH	79.96	64.20	44.94	100 (19)[a]	50.18
mPPE-CH_2OH	40.80	83.60	52.30	46.90	100 (7)[a]
mPPE-OCH_3	100 (10)[a]	100 (13)[a]	100 (3)[a]	100 (9)[a]	88.08
mPPE-H	7.46	-	8.98	-	8.56

[a] In these simulations, some water molecules remained in the channel for the duration of the 100 ns simulation; the number of molecules doing so is listed in parentheses.

From simulations involving the moderately hydrophilic channel, (mPPE-OCH_3), we observed that water transport through the channel was either small (simulations with the SPC, TIP3P, TIP4P and TIPP water models) or nonexistent (simulations with SPC/E water model). This low level of diffusion was observed even though simulations with all five models showed that water was able to enter the channel. A more detailed analysis revealed that water molecules tended to diffuse in and move back out of the channel ends, while the middle section of the channel was not completely filled with water molecules. The diffusivity of the water inside mPPE-OCH_3 channels is comparable to that observed for diffusion in the hydrophilic channels (see Table 6.4).

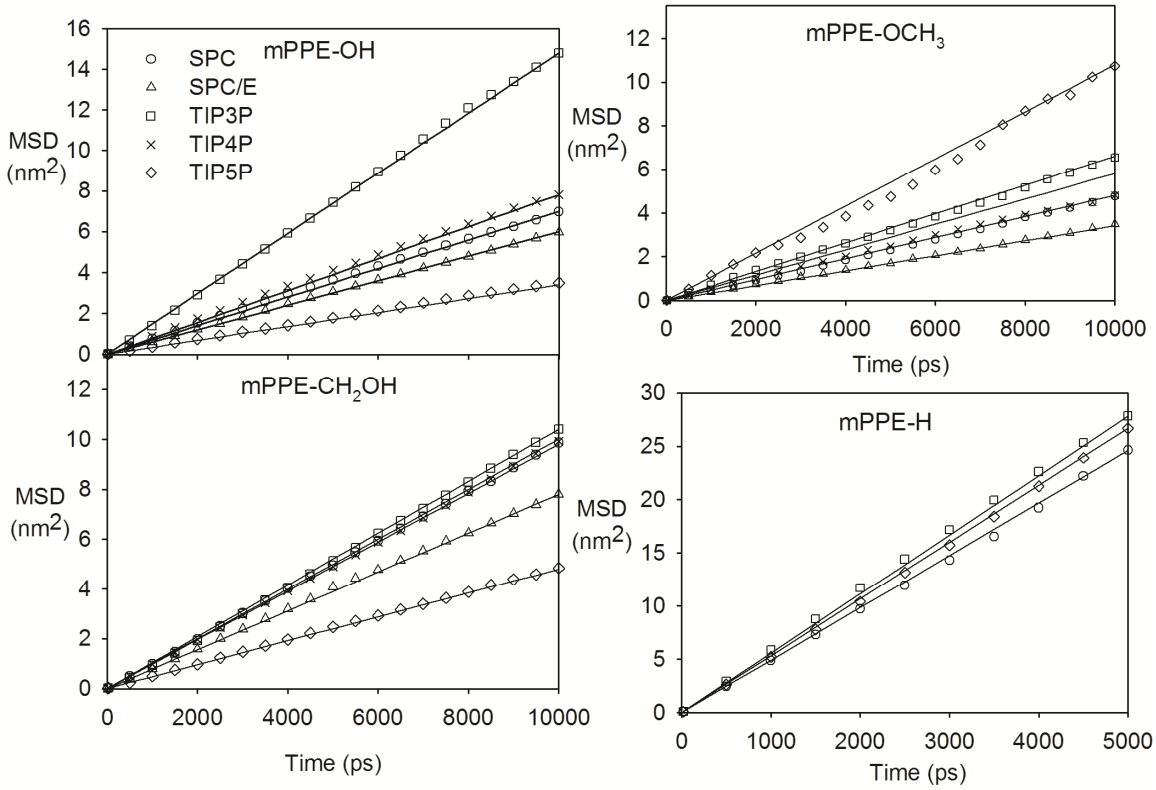


Figure 6.8 1-D mean square displacement $(MSD)_z$ in the z-direction of water molecules inside mPPE nanochannels (mPPE-OH, mPPE-OCH_3, mPPE-CH_2OH and mPPE-H) over the simulated time, for the five studied water models: SPC, SPC/E, TIP3P, TIP4P and TIP5P. The solid lines represent trends predicted from the diffusivities reported in Table 6.4.

As noted in the previous section, simulation results varied with respect to the water model used. In simulations of mPPE-OH/TIP3P channels, we observed a significantly higher number of water molecules entering and diffusing through the channel, compared to simulation with the other water models. Also, the water diffusivity inside the mPPE-OH channel that was predicted from simulations using the TIP3P water model was higher

than those of other water models. This is in agreement with prior modeling results, which showed that the TIP3P water model yields overpredicts the magnitude of the diffusivity of bulk water (see Table 6.4). However, a similar result was not observed in simulations with the mPPE-CH_2OH channel. With this functionalized channel, all five water models showed a similar number of molecules diffusing in and through the channels, with similar diffusivities. It is possible that the smaller effective diameter of the mPPE-CH_2OH nanochannel (0.396 nm, compared to 0.58 nm effective diameter of the mPPE-OH nanochannel) allows for greater interaction between water and the *endohelix* functional groups, thereby reducing the number of interactions among water molecules inside the channels. The higher effective diameter of the mPPE-OH channel allows more interaction among water molecules, resulting in more bulk-like behavior (i.e., the water is more mobile and thus has a higher diffusivity), explaining why the trend among water model diffusivities in mPPE-OH resembles that for simulations of bulk water.

The simulation results indicate that the selectivity for water of the mPPE channels could be controlled by using appropriate *endohelix* functional groups. For example, different functional groups may be used to block (mPPE-CH_3) or allow (mPPE-OH, mPPE-CH_2OH and mPPE-OCH_3) water to cross the channel, and to vary the exchange rate of water from one side of the mPPE channel to the other (mPPE-OH, mPPE-CH_2OH and mPPE-OCH_3). They also provide evidence that, in a similar manner, it may be possible to achieve specific mPPE channel selectivity towards a range of chemicals other than water.

6.3.4 Effect of water models on water diffusion simulation in nano systems

The observed variations in simulation results for different water models are mainly qualitative, and in general, the differences in molecular dynamics model parameters for water yield different predicted values for the parameters of interest (e.g., water diffusivity). However, in simulations with the mPPE-H channels, there are significant quantitative differences between the simulation results. As presented earlier for this system, simulation with the SPC, TIP3P and TIP5P water models predicted that water would readily diffuse into the channel and form stable water columns, while simulations employing the SPC/E and TIP4P water models yielded systems that were unable to maintain a continuous water column inside the mPPE-H channel.

The anomalies in water diffusion behavior can be attributed to differences in the water models. For example, the SPC and SPC/E water models, both model water using a three-atom centered interaction sites, but the partial charges assigned to the oxygen and hydrogen atoms differ for the two models. In the SPC/E model, the partial charges are set higher to account for the polarizing effect of water networks; however, in the present simulation study, this small change seems to have a significant effect on the diffusion of water through the hydrophobic channels. To illustrate this, five additional simulations were conducted with mPPE-H channels and SPC type water models, in which the partial charges of oxygen and hydrogen were varied between the values used for SPC and SPC/E models. At the start of these simulations, there were no water molecules inside the channels and short (36 ns) simulations were conducted to investigate whether water

molecules could diffuse in and form a stable water column inside an mPPE-H channel. The results summarized in Table 6.6 indicate that the observed differences in diffusion behaviors are attributable to variations in the interatomic interaction parameters for the specific water models. Though not completely surprising given the many ongoing efforts to develop molecular models for water, these diffusion simulations illustrate the importance of using several models for water to try and ascertain what the true diffusion behavior would be for water molecules in hydrophobic mPPE channels.

Table 6.6 Effect of water partial atomic charges on the stability of the water column inside mPPE-H nanochannels for the SPC and SPC/E water models.

Model	Oxygen partial charge (q_0)	Hydrogen partial charge (q_1)	Simulation observations
SPC[a]	-0.8200	0.41000	Filled pore[b]
SPC-1	-0.8269	0.41435	Filled pore
SPC-2	-0.8292	0.41460	Filled pore
SPC-3	-0.8315	0.41575	Not filled pore[c]
SPC-4	-0.8338	0.41690	Not filled pore
SPC-5	-0.8407	0.42035	Not filled pore
SPC/E[a]	-0.8476	0.42380	Not filled pore

[a] Oxygen and hydrogen partial atomic charges of water (as reported in the Ref. 61 and 62).
[b] The water column formed then stayed inside the mPPE-H channel for the length of the simulation.
[c] The water column disconnected and water molecules withdraws from the inside of the mPPE-H channel.

6.4 Conclusions

In this chapter, we reported the simulation results for water diffusion through *endohelix* functionalized mPPE nanochannels. The results show that all five of the functionalized mPPEs studied were able to form stable helical channels that penetrated completely through an octane monolayer, connecting two reservoirs of water that were otherwise separated by the monolayer. Such results indicate the potential of using mPPE helical structures as functionalized nanochannels to selectively conduct the flow of water or other small molecules through otherwise impermeable structures. We also showed that mPPEs are versatile porous nanostructures for membrane, sensor, and cellular trafficking type applications, and that simple variations of the polymer chemical functionality can significantly alter the diffusion characteristics for water through the channel. By using the appropriate *endohelix* functional groups, the mPPE channel interior environment properties (e.g., polarity or interior diameter) can be selectively manipulated for a given application. This concept was demonstrated in our simulation results in which the movement of water through a given mPPE channel could be selectively blocked or permitted depending on the *endohelix* functional groups of the mPPE, and that the flow rates of water through such channels could similarly be controlled. The hydrophilic mPPE channels having polar *endohelix* functional groups exhibit similarities to biological channels, and the hydrophobic mPPE channel having a large diameter (mPPE-H) behaves similar to carbon nanotube channels. These simulation results indicate that mPPE channels could be used for a wide range of diffusion related applications, and suggest that such systems should be explored experimentally.

6.5 References

1. Hummer, G.; Rasaiah, J. C.; Noworyta, J. P., *Nature* **2001,** *414* (6860), 188-190.

2. Kalra, A.; Garde, S.; Hummer, G., *Proc. Natl. Acad. Sci. U.S.A.* **2003,** *100* (18), 10175-10180.

3. Lehninger, A. L.; Nelson, D. L.; Cox, M. M., *Lehninger principles of biochemistry*. 5th ed.; W.H. Freeman: New York, **2008**; pp 371-413.

4. Liu, B.; Li, X. Y.; Li, B. L.; Xu, B. Q.; Zhao, Y. L., *Nano Lett.* **2009,** *9* (4), 1386-1394.

5. Luckey, M., *Membrane structural biology : with biochemical and biophysical foundations*. Cambridge University Press: Cambridge ; New York, **2008**; pp 13-26, 138-139, 196-202, 213-241.

6. Mukherjee, B.; Maiti, P. K.; Dasgupta, C.; Sood, A. K., *J. Nanosci. Nanotechnol.* **2007,** *7* (6), 1796-1799.

7. Rivera, J. L.; Starr, F. W., *J. Phys. Chem. C* **2010,** *114* (9), 3737-3742.

8. Tu, Y. S.; Xiu, P.; Wan, R. Z.; Hu, J.; Zhou, R. H.; Fang, H. P., *Proc. Natl. Acad. Sci. U.S.A.* **2009,** *106* (43), 18120-18124.

9. van Hijkoop, V. J.; Dammers, A. J.; Malek, K.; Coppens, M. O., *J. Chem. Phys.* **2007,** *127* (8), 085101(1-10).

10. Zhu, F. Q.; Tajkhorshid, E.; Schulten, K., *Biophys. J.* **2004,** *86* (1), 50-57.

11. Zhong, Q. F.; Jiang, Q.; Moore, P. B.; Newns, D. M.; Klein, M. L., *Biophys. J.* **1998,** *74* (1), 3-10.

12. Zhong, Q.; Moore, P. B.; Newns, D. M.; Klein, M. L., *Febs Letters* **1998,** *427* (2), 267-270.

13. Kaucher, M. S.; Peterca, M.; Dulcey, A. E.; Kim, A. J.; Vinogradov, S. A.; Hammer, D. A.; Heiney, P. A.; Percec, V., *J. Am. Chem. Soc.* **2007,** *129* (38), 11698-11699.

14. Hill, D. J.; Mio, M. J.; Prince, R. B.; Hughes, T. S.; Moore, J. S., *Chem. Rev.* **2001,** *101* (12), 3893-4011.

15. Lahiri, S.; Thompson, J. L.; Moore, J. S., *J. Am. Chem. Soc.* **2000,** *122* (46), 11315-11319.

16. Nelson, J. C.; Saven, J. G.; Moore, J. S.; Wolynes, P. G., *Science* **1997,** *277* (5333), 1793-1796.

17. Nguyen, H. H.; McAliley, J. H.; Batson, W. A.; Bruce, D. A., *Macromolecules* **2010,** *43* (14), 5932-5942.

18. Brunsveld, L.; Prince, R. B.; Meijer, E. W.; Moore, J. S., *Org. Lett.* **2000,** *2* (11), 1525-1528.

19. Cary, J. M.; Moore, J. S., *Org. Lett.* **2002,** *4* (26), 4663-4666.

20. Nguyen, H. H.; McAliley, J. H.; Bruce, D. A., *J. Polymer Sci. Polymer Chem.* **2011,** *in press*.

21. Nguyen, H. H.; McAliley, J. H.; Bruce, D. A., *Macromolecules* **2011,** *44* (1), 60-67.

22. Prince, R. B.; Saven, J. G.; Wolynes, P. G.; Moore, J. S., *J. Am. Chem. Soc.* **1999,** *121* (13), 3114-3121.

23. Prince, R. B.; Okada, T.; Moore, J. S., *Angew. Chem. Int. Ed.* **1999,** *38* (1-2), 233-236.

24. Stone, M. T.; Fox, J. M.; Moore, J. S., *Org. Lett.* **2004,** *6* (19), 3317-3320.

25. Stone, M. T.; Heemstra, J. M.; Moore, J. S., *Acc. Chem. Res.* **2006,** *39* (1), 11-20.

26. Stone, M. T.; Moore, J. S., *Org. Lett.* **2004,** *6* (4), 469-472.

27. Tan, C. Y.; Pinto, M. R.; Kose, M. E.; Ghiviriga, I.; Schanze, K. S., *Adv. Mater.* **2004,** *16* (14), 1208-1212.

28. Hecht, S.; Khan, A., *Angew. Chem. Int. Ed.* **2003,** *42* (48), 6021-6024.

29. Adisa, B.; Bruce, D. A., *J. Phys. Chem. B* **2005,** *109* (15), 7548-7556.

30. Adisa, B.; Bruce, D. A., *J. Phys. Chem. B* **2005,** *109* (42), 19952-19959.

31. Adisa, B. Synthesis, characterization and modeling of novel polymers and polymer-templated mesoporous materials. Ph.D. Dissertation, Clemson University, Clemson, SC, U.S.A., **2005**.

32. Elmer, S. P.; Pande, V. S., *J. Chem. Phys.* **2004,** *121* (24), 12760-12771.

33. Elmer, S. P.; Park, S.; Pande, V. S., *J. Chem. Phys.* **2005,** *123* (11), 114903(0-7).

34. Elmer, S. P.; Park, S.; Pande, V. S., *J. Chem. Phys.* **2005,** *123* (11), 114902(0-14).

35. *High performance computing*. http:/citi.clemson.edu/hpc

36. Berendsen, H. J. C.; van de Spoel, D.; van Drunen, R., *Comp. Phys. Comm.* **1995,** *91*, 43-56.

37. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E., *J. Chem.Theory Comput.* **2008,** *4* (3), 435-447.

38. Lindahl, E.; Hess, B.; van de Spoel, D., *J. Mol. Mod.* **2001,** *7*, 306-317.

39. van der Spoel, D.; E., L.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C., *J. Comp. Chem.* **2005,** *26*, 1701-1718.

40. van der Spoel, D.; Lindahl, E.; Hess, B.; van Buuren, A. R.; Apol, E.; Meulenhoff, P. J.; Tieleman, D. P.; Sijbers, A. L. T. M.; Feenstra, K. A.; van Drunen, R.; Berendsen, H. J. C. GROMACS USER MANUAL version 4.5.4 www.gromacs.org/documentation (accessed on 08/03/2011).

41. Jorgensen, W. L.; Madura, J. D.; Swenson, C. J., *J. Am. Chem. Soc.* **1984,** *106* (22), 6638-6646.

42. Jorgensen, W. L., *J. Phys. Chem.* **1986,** *90* (7), 1276-1284.

43. Jorgensen, W. L.; Maxwell, D. S.; TiradoRives, J., *J. Am. Chem. Soc.* **1996,** *118* (45), 11225-11236.

44. Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L., *J. Phys. Chem. B* **2001,** *105* (28), 6474-6487.

45. Price, M. L. P.; Ostrovsky, D.; Jorgensen, W. L., *J. Comput. Chem.* **2001,** *22* (13), 1340-1352.

46. Rizzo, R. C.; Jorgensen, W. L., *J. Am. Chem. Soc.* **1999,** *121* (20), 4827-4836.

47. Watkins, E. K.; Jorgensen, W. L., *J. Phys. Chem. A* **2001,** *105* (16), 4118-4125.

48. Humphrey, W.; Dalke, A.; Schulten, K., *J. Molec. Graphics.* **1996,** *14* (1), 33-38.

49. Roux, B.; Karplus, M., *Biophys. J.* **1991,** *59* (5), 961-981.

50. Sagnella, D. E.; Voth, G. A., *Biophys. J.* **1996,** *70* (5), 2043-2051.

51. Kozono, D.; Yasui, M.; King, L. S.; Agre, P., *J. Clin. Invest.* **2002,** *109* (11), 1395-1399.

52. *Materials Studio*, 5.0; Accelrys Software Inc: San Diago, CA, **2008**.

53. Weber, M. E.; Schlesinger, P. H.; Gokel, G. W., *J. Am. Chem. Soc.* **2005,** *127* (2), 636-642.

54. Zhu, C. Y.; Byrd, R. H.; Lu, P. H.; Nocedal, J., *Acm T. Math. Software* **1997,** *23* (4), 550-560.

55. Bussi, G.; Donadio, D.; Parrinello, M., *J. Chem. Phys.* **2007,** *126*(1), 014101(0-7).

56. Parrinello, M.; Rahman, A., *J. Appl. Phys.* **1981,** *52* (12), 7182-7190.

57. Lide, D. R.; Kehiaian, H. V., *CRC handbook of thermophysical and thermochemical data*. CRC Press: Boca Raton, **1994**; pp 518.

58. Darden, T.; York, D.; Pedersen, L., *J. Chem. Phys.* **1993,** *98* (12), 10089-10092.

59. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., *J. Chem. Phys.* **1995,** *103* (19), 8577-8593.

60. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J., Interaction models for water in relation to protein hydration. In *Intermolecular forces*, Pullman, B., Ed. Reidel Dordrecht Publishing Company: **1981**; pp 331-342.

61. Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P., *J. Phys. Chem.* **1987,** *91* (24), 6269-6271.

62. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., *J. Chem. Phys.* **1983,** *79* (2), 926-935.

63. Mahoney, M. W.; Jorgensen, W. L., *J. Chem. Phys.* **2000,** *112* (20), 8910-8922.

64. Bondi, A., *J. Phys. Chem.* **1964,** *68* (3), 441-451.

65. Hill, D. J.; Moore, J. S., *Proc Natl Acad Sci U S A* **2002,** *99* (8), 5053-7.

66. Mahoney, M. W.; Jorgensen, W. L., *J. Chem. Phys.* **2001,** *114* (1), 363-366.

67. van der Spoel, D.; van Maaren, P. J.; Berendsen, H. J. C., *J. Chem. Phys.* **1998,** *108* (24), 10220-10230.

68. Eisenberg, D.; Kauzmann, W., *The structure and properties of water*. Oxford University Press: Oxford, **1969**; pp 217-222.

CHAPTER 7

CONCLUSIONS AND RECOMMENDATIONS

The results described herein have examined the use of functionalized *meta*-poly(phenylene ethynylene) (mPPE) helical structures for two new applications: as chiral and achiral supports for transition metal catalysts and secondly, as nanochannels for selectively facilitating the transport of water and other small molecules Additionally, this study examined the applicability of molecular modeling as a guide for the synthesis of these unique polymer structures. Achieving this goal required a combination of experimental and modeling efforts, which included: i) the use of molecular dynamic (MD) simulations to predict the folding behaviors of functionalized mPPEs and determine their applicability for applications requiring ordered helical polymer structures; and ii) the use of molecular simulation results to guide the synthesis and characterization of novel functionalized mPPEs.

CONCLUSIONS

1. The Replica Exchange MD (REMD) simulation procedure, which employed the GROMACS simulation program and OPLS force field, provided reliable predictions for the folding behaviors of functionalized mPPEs in a range of solvent conditions. The typical time need for one REMD simulation, which used 40 of the previously described

CPUs is two days. This procedure was validated against experimental data for ester functionalized mPPEs. In all cases, the simulation results were in excellent agreement with all available experimental data.

2. A combinatorial study examining the effects of functional groups and solvent conditions on mPPE folding behaviors demonstrated that simple structural relationships could not accurately predict the folding behaviors of an mPPE given its functional groups and solvent conditions, implying that molecular modeling was both an effective and essential tool for predicting polymer secondary structure formation.

3. REMD results showed that having an alternating arrangement of ester and other *exohelix* functional groups on the mPPE backbone was one way to produce functionalizable and foldable mPPE structures. The second advantage of this structure is the enhanced mPPE solubility provided by the ether pendants attached to the ester *exohelix* functional groups. Two new mPPEs having both ester and nitrile *exohelix* functional groups were synthesized so as to validate the REMD simulation procedure.

4. Five other new functionalized mPPEs having an alternating arrangement of ester and other *exohelix* functional groups, as described in Conclusion 3, were synthesized and their folding behaviors were characterized using UV absorbance and fluorescence emission spectra.

5. Two new mPPEs having both ester and amine *exohelix* functional groups were found to be fully soluble in water and alcohol, and both folded into stable helical conformations in water.

6. A manganese(salen) mPPE complex (Mn(salen)-mPPE) was successfully synthesized using imine functionalized mPPE helical structures as the salen ligand. This manganese complex was found to have some catalytic activity for styrene epoxidation reactions.

7. REMD simulation results of mPPEs provided evidence that having hydrogen bonds among *endohelix* functional groups can strongly increase the stability of helical structures, allowing the possibility of synthesizing customized *exohelix* functionalized mPPE-helical materials.

8. REMD simulation results provided evidence that the ester *exohelix* functionalized mPPE can form stable helical conformations with many different *endohelix* functional groups.

9. The concept of using functionalized helical mPPEs as functionalizable nanochannels for water was demonstrated via MD simulations. Five mPPE nano channels consisting of 96 monomers were found to be stable in octane – water systems during 100 ns simulations.

10. Results from MD simulations of water diffusion through mPPE nanochannels indicated that water could rapidly diffuse through the channel having simple hydrogen *endohelix* functional groups but not diffuse into channels having methyl *endohelix* functional groups. Water was also able to diffuse through channels having polar or moderately polar functional groups. The results also showed that diffusion simulations were very sensitive to the specific water model parameters employed in the simulations.

RECOMMENDATIONS

Based on the works conducted for this study, the following research would be recommended:

1. The REMD procedure can be expanded to study chain length effects on the folding behaviors of functionalized mPPEs. Evidence from synthesized mPPEs and related simulation results indicate that mPPEs having more than 12 aromatic rings can form stable helical structures, as in Chapter 3, especially for mPPEs whose 12-mer structures were categorized as partially folding in acetonitrile. To this end, we have started a molecular simulation effort examining several mPPEs having chain lengths ranging from 14-mer up to 24-mer in acetonitrile to examine the effect of chain length on the folding behaviors of functionalized mPPEs.

2. The REMD procedure can also be expanded to include many more functional groups (such as sulfo or fluorinated functional groups) and solvents conditions (such as fluorinated solvents or dense CO_2 modified solvents), as well as to study other factors affecting the mPPE folding behaviors, such as polymer concentration and temperature.

3. The modeling system used to study the diffusion of water through mPPE nanochannels can be used to simulate the diffusion of different small molecules, such as low molecular alcohols, hydrocarbons, and their mixtures, with a goal of exploring the use of mPPE nanochannels as selective channels for membrane filtration applications.

4. In Chapter 4, the host-guest interactions between the mPPE helical conformation and the chiral guest, and the chiral pendants attached to the polymer back bone lead to the formation of mPPE chiral helical structures. Those interactions suggest modeling studies using MD and/or quantum mechanics simulations to investigate this phenomenon. Results from simulation studies can provide important information that is difficult to access through experiments, for example, the direction of the helix coil or the strength of the binding forces.

5. Because the synthesized functionalized mPPEs in this study have broad molecular weight distributions, assessment of the relationship between the chain lengths of those polymers and their folding behaviors are only qualitative. An experimental study using preparative GPC is suggested to collect mPPE fractions having narrow molecular weight distributions, thus, enabling one to quantitatively evaluate that relationship.

6. The difficulty of using polystyrene with narrow molecular weight distribution standards in Gel Permeation Chromatography (GPC) analysis of mPPE chain length suggests the need for a systematic study requiring the synthesis of precise chain length mPPEs either to use as standards or to establish a calibration curve, between the chain length of the mPPEs and the molecular weight of polystyrene with narrow molecular weight distribution standards. The results from this study will be useful for the study described in Recommendation 5. The molecular weight distributions of mPPE samples could also be estimated using GPC equipped with a light scaterring detector and/or an instrinsic viscosity detector.

7. The synthesis route to produce the Mn(salen)-mPPE complex using imine functionalized mPPE helical structures as the salen ligand in this study could be employed to synthesize other metal salen complexes, such as Cobalt.

8. As proposed in Chapter 4, the successful formation of the Mn(salen)-mPPE complex suggests the possibility of synthesizing a chiral type of this complex with chiral mPPE helical conformations as the salen ligand. Subsequent studies could employ either a chiral guest or chiral pendants approach to form the imine functionalized mPPE chiral helical structures.

9. The low reaction yield of the styrene epoxidation reaction using the Mn(salen)-mPPE complex as catalyst suggests an investigation on: i) the Mn(salen)-mPPE complex formation conditions, such as solvent, reaction time, temperature, etc., to increase the Manganese loading; and ii) the epoxidation reaction conditions including, substrates (other olefins, such as 1-hexene), the choice of oxidizing agents (non-chlorinated system such as *meta*-chloroperoxybenzoic acid), temperature, reaction time, etc.

10. Results from simulations provided predictions that mPPEs having hydrogen bonding between *endohelix* functional groups are capable of forming stable helical conformations, allowing the possibility of customized *exohelix* functionalized mPPE helical structures. An experimental study is suggested to synthesize and characterize examples of these functionalized mPPEs; for example, prepare and test an mPPE having hydroxyl *endohelix* functional groups.

APPENDIX A

PREPARING A PDB FILE

The pdb file is the file containing the positions of all atoms in the simulated system. This is the required input information for a molecular dynamics simulation. This section presents the procedure to build repeat units and polymers using Materials Studio software, to export the pdb file, and to modify the pdb file to Gromacs format.

A.1 Build the ester functionalized mPPE structure using Materials Studio

1) Open Materials Studio software.

2) Select new project. Name the new project. Materials Studio windows will appear like in Figure A.1.

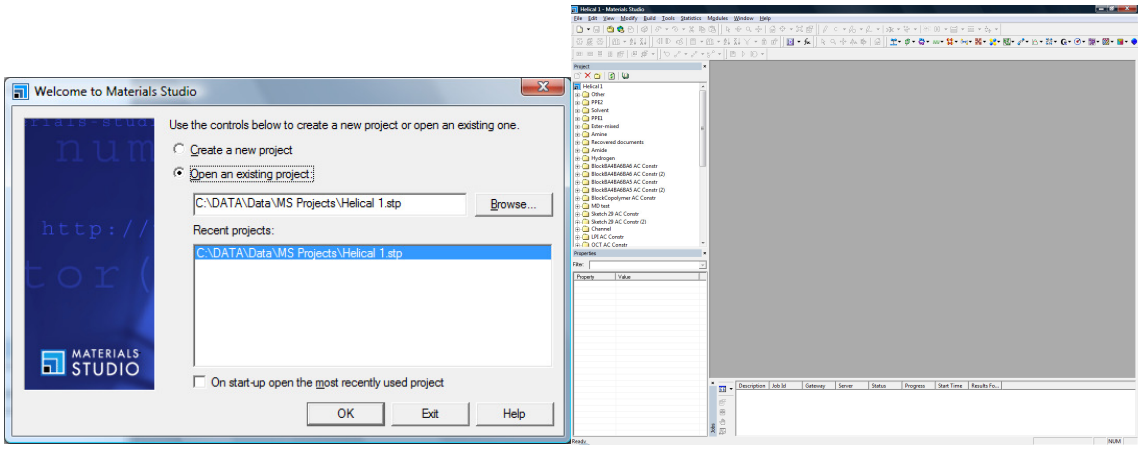


Figure A.1 Starting windows of Materials Studio.

3) Draw the repeat unit

- Click on **New** at the left conner, under the **File** menu. Select **3D atomistic document** from the drop-down list.

- Follow the instructions on the Help page to draw the repeat unit. Figure A.2 represents a repeat unit as finished for the ester functionalized mPPE.

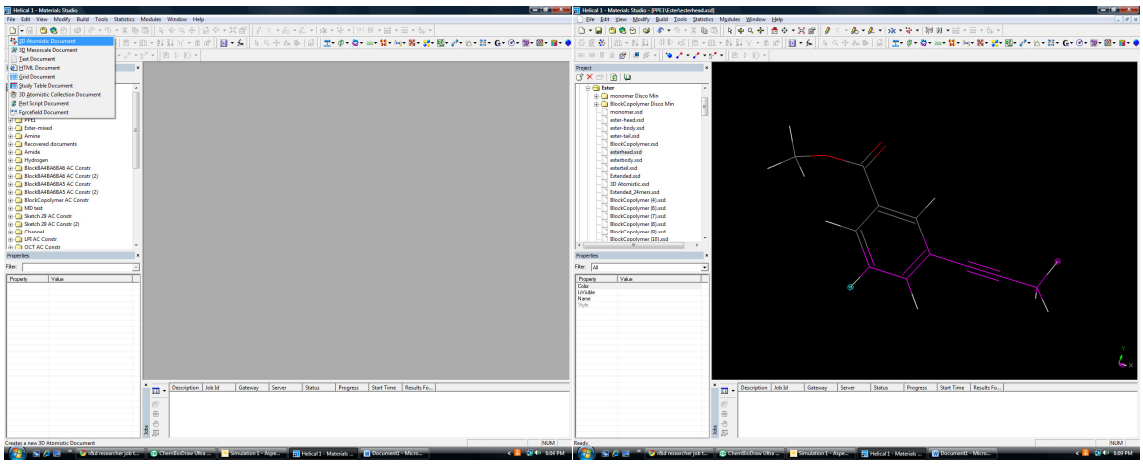


Figure A.2 New structure menu and representative mPPE repeat unit.

4) Define the repeat unit

- Click on **Build** menu, select **Build polymer**, then **Repeat Unit**. A small window will appear. Select the head atom of the the repeat unit, then click **Head atom** on the window. Do the same for the **Tail atom**. On the Figure A.3, the back bone of the repeat unit is in purple color.

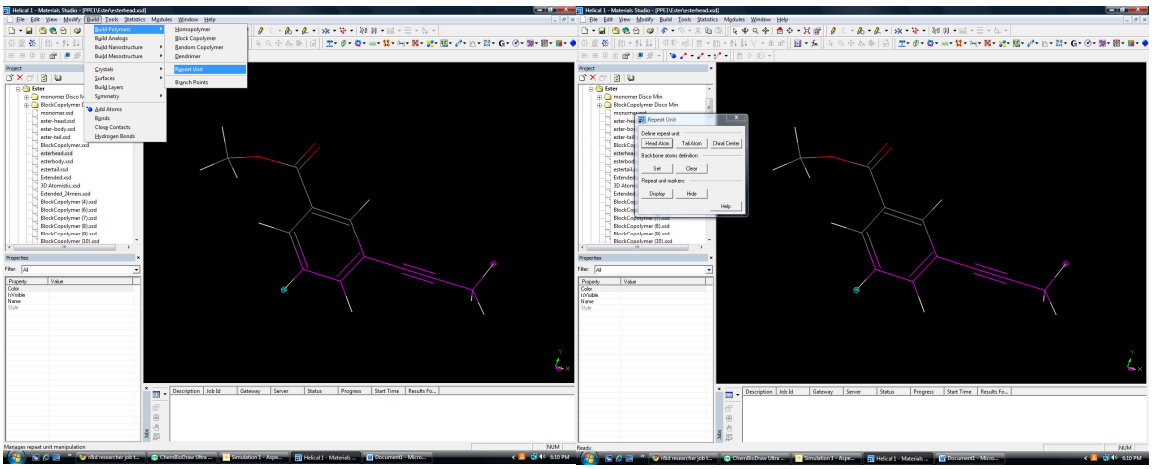


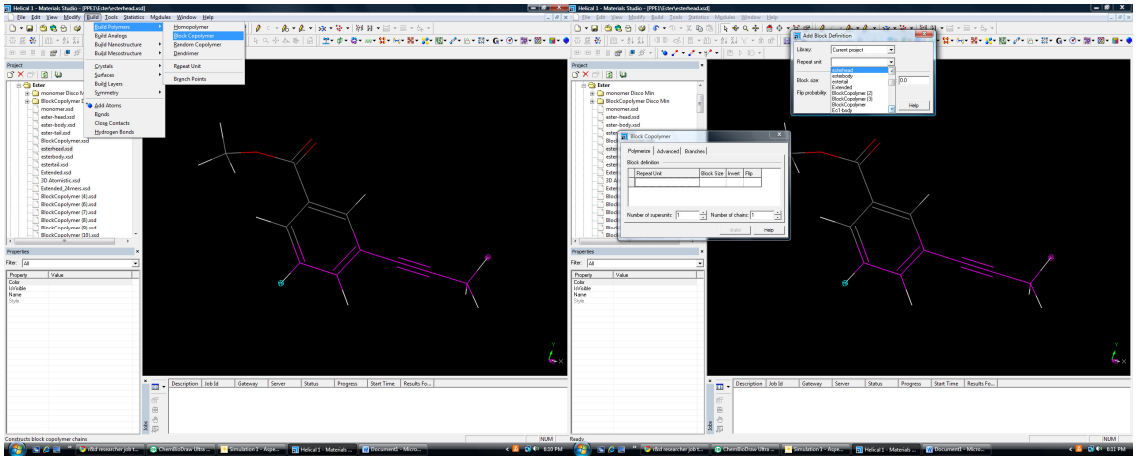
Figure A.3 Repeat unit definition menu.

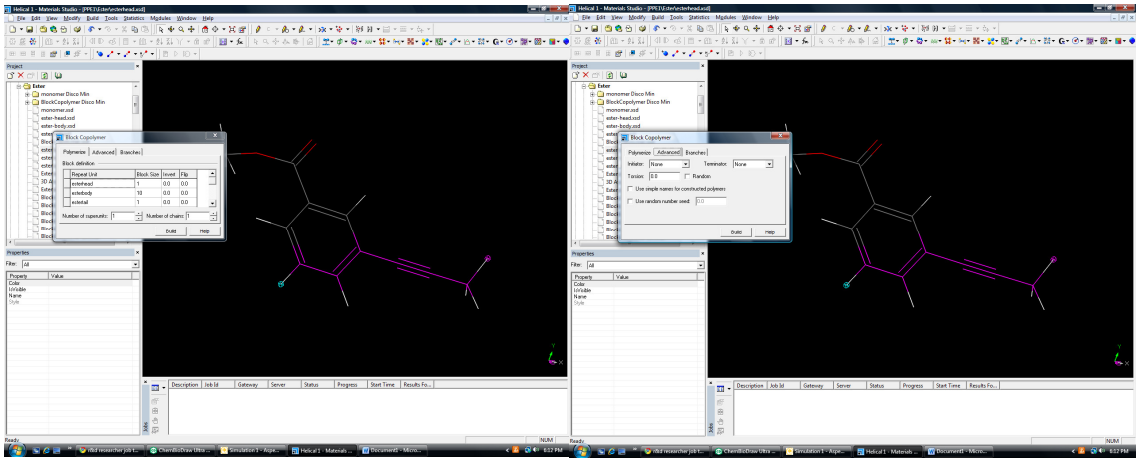
5) Build the polymer using Polymer builder

- Click on the **Build** menu, select **Build polymer**, then **Block copolymer**. A small window will appear.

- Select the name of the **Repeat unit** and enter the number appearing in the polymer. For each mPPE, three repeat units, each for the head, body and tail sections, are needed to be selected.

The procedure is presented on Figure A.4.





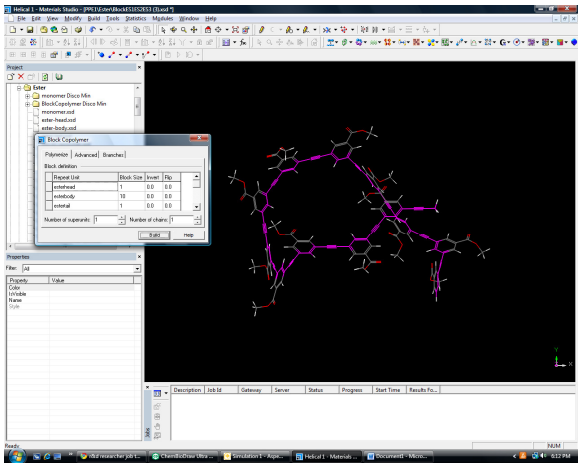


Figure A.4 Illustration of polymer building procedure.

6) Export the pdb file

-Select the **File** menu, then export. A new window will appear; select the **pdb** file type from the drop-down list (as shown in Figure A.5).

- Enter the file name and then click **Export.**

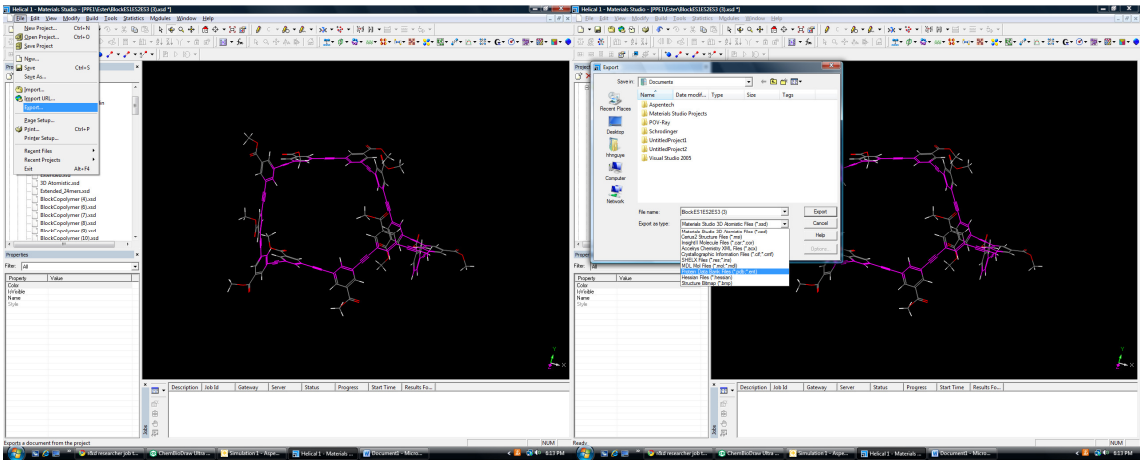


Figure A.5 Illustration of pdb files exporting procedure.

A.2 Modify the pdb file from Materials Studio to Gromacs format

The pdb file exported from Materials Studio is not compatible with Gromacs. The first two lines need to be removed. Further, Materials Studio considers the polymer as 1 residue whereas Gromacs considers the polymer as multiple residues as defined in section D.2. Thus, the residue number needs to be changed. For more information about the format of the pdb files, see Gromacs documentation at www.gromacs.org/documentation.

An example pdb file exported from Materials Studio:

```
REMARK   Materials Studio PDB file
REMARK   Created:  Thu Jun 26 12:54:11 Eastern Standard Time 2008
ATOM     1  H1  ES1 B   0     -15.037  -9.964   1.061  1.00  0.00     H
ATOM     2  C4  ES1 B   0     -14.469  -9.670   1.961  1.00  0.00     C
ATOM     3  C5  ES1 B   0     -13.676  -8.557   1.689  1.00  0.00     C
……..……..

ATOM   472  H43 ES3 B   0      31.008  36.773  28.577  1.00  0.00     H
ATOM   473  H42 ES3 B   0      32.515  35.910  28.176  1.00  0.00     H
TER
```

The pdb file used by Gromacs:

```
ATOM     1  H1  ES1 B   1     -15.037  -9.964   1.061  1.00  0.00     H
ATOM     2  C4  ES1 B   1     -14.469  -9.670   1.961  1.00  0.00     C
ATOM     3  C5  ES1 B   1     -13.676  -8.557   1.689  1.00  0.00     C
……..……..

ATOM   472  H43 ES3 B  12      31.008  36.773  28.577  1.00  0.00     H
ATOM   473  H42 ES3 B  12      32.515  35.910  28.176  1.00  0.00     H
TER
```

The following python script is to change the residue number:

```
#Reformat file *.pdb, change the residue I.D.
#The first two lines of the pdb file need to be reomove before using
#this script to change the residue I.D.
#Change Residual sequence number
#Total '++MAX++' residues, the first residue has '++ATOM++' atoms
```

```
#Change ‘++FILENAME++’ to the name of the pdb file from Materials
Studio
#change '++ATOM++' to the number of atoms of the head residues in your
polymer
#change '++MAX++' to the number of residues in your polymer


a=++ATOM++
b=a-3
c=b+1
d=a-1
e=b-1

max=++MAX++
f=open("++FILENAME++.pdb","r+")
f.seek(0)
i=1
j=0
while j<a:
      f.seek(84*j+25)
      s=str(i)
      f.write(s)
      j=j+1
i=2
j=1
if max < 10:
      while i<max:
            while j<b:
                  f.seek(84*(d+j+e*(i-2))+25)
                  k=i
                  s=str(k)
                  f.write(s)
                  j=j+1
            i=i+1
            j=1
      i=max
      j=1
      while j<c:
            f.seek(84*(d+j+e*(i-2))+24)
            s=str(i)
            f.write(s)
            j=j+1
elif:
      while i<10:
                while j<b:
                        f.seek(84*(d+j+e*(i-2))+25)
                        k=i
                        s=str(k)
                        f.write(s)
                        j=j+1
                i=i+1
                j=1
      i=10
      j=1
```

```
      while i<max:
            while j<b:
                  f.seek(84*(d+j+e*(i-2))+24)
                  k=i
                  s=str(k)
                  f.write(s)
                  j=j+1
            i=i+1
            j=1
      i=max
      j=1
      while j<c:
            f.seek(84*(d+j+e*(i-2))+24)
            s=str(i)
            f.write(s)
            j=j+1
f.close()
```

The bash file to execute the above python file:

```
#!/bin/bash
echo 'Name of the .pdb file from Materials Studio (‘temp’ in
temp.pdb)?'
read FILENAME
echo “The new file will be name ${FILENAME}_formatted.pdb”
echo ‘Number of atoms in the first residue of the mPPE chain?’
read ATOM
echo ‘Number of residues in the mPPE chain?’
read RESIDUE
echo ‘name of the python pdb formatting script?’
read PYTHONNAME
cat | grep ‘ATOM’ ${FILENAME}.pdb > 1.pdb
sed –e “s/++FILENAME++/1/g” ${PYHTONNAME} > 1
sed –e “s/++ATOM++/${ATOM}/g” 1 > 2
sed –e “s/++MAX++/${RESIDUE}/g” 2 > 3
python 3
rm 1 2 3
editconfig –f 1.pdb –o ${FILENAME}_reformated.pdb
rm 1.pdb
```

New version of the reformatting code for multiple residues having the same number of atoms:

```
#reformat file *.pdb
#directory /root/desktop/Python/
import math

max_res=RESIDUE_NUMBER
a=ATOM_NUMBER
f=open("FILE_NAME","r+")
f.seek(0)
j=0
i=1
while i<(max_res+1):
      indent = 25-int(math.log10(i))
      while j<a:
                f.seek(84*(j+(i-1)*a)+indent)
                k=i
                s=str(k)
                f.write(s)
                j=j+1
        i=i+1
        j=0
f.close()
```

New version of the python reformatting code:

```
#Reformat file *.pdb, change the residue I.D.
#The first two lines of the pdb file need to be reomove before using
#this script to change the residue I.D.
#Change Residual sequence number
#Total 'RESIDUE' residues, the first residue has 'ATOM' atoms

import math
FILENAME =raw_input('Enter file name (*.pdb):')
ATOM = None
while not ATOM:
      try:
            ATOM=int(raw_input('Enter the number of atoms in the first
      residue:')
      Except ValueError:
            print 'Invalid number')

RESIDUE = None
while not RESIDUE:
      try:
```

```
            RESIDUE=int(raw_input('Enter the number of residue:')
      Except ValueError:
            print 'Invalid number')

a=ATOM
b=a-3
c=b+1
d=a-1
e=b-1

max=RESIDUE
f=open(FILENAME,"r+")
f.seek(0)
i=1
j=0
while j<a:
      f.seek(84*j+25)
      s=str(i)
      f.write(s)
      j=j+1
i=2
j=1
while i<(max_res):
      indent = 25-int(math.log10(i))
            while j<b:
                  f.seek(84*(d+j+e*(i-2))+indent)
                  k=i
                  s=str(k)
                  f.write(s)
                  j=j+1
            i=i+1
            j=1
i=max
j=1
while j<c:
      f.seek(84*(d+j+e*(i-2))+24)
      s=str(i)
      f.write(s)
      j=j+1
f.close()
```

APPENDIX B

PREPARING mPPE SIMULATION SYSTEMS

B.1 Preparation of the simulation system

After reformatting the pdb file, the next step is to prepare the simulation system. This step includes the following sub-steps:

1. Import the pdb file into gromacs using the *pdb2gmx* command. This command will import data from the pdb file, and use information from the force field to generate several files. The three most important files are: *.gro, *.top and *.itp

2. Define the simulation box using the *editconf* command. This command will define the simulation box, size, position of the polymer inside the box.

3. Conduct an energy minimization simulation for the ester functionalized mPPE structure in vacuum using the *mdrun* command. The command *grompp* is used to generate the file *.tpr required by *mdrun*.

4. Solvate the simulation box with solvent molecules using the *genbox* command. Two files are needed for each solvent type: *.itp and *.gro.

All the structural information of a solvent model is listed in an *.itp file. For example, the *.itp file for acetonitrile:

```
[ moleculetype ]
; name  nrexcl
AN1   3

 [ atoms ]
;   nr    type   resnr  residu  atom    cgnr    charge
      mass
1     opls_094    1     AN1   N     1     -0.430      14.00670
2     opls_095    1     AN1   C1    1      0.280      12.01100
3     opls_096    1     AN1   C2    1      0.150      15.03500


[ bonds ]
;  ai    aj funct          b0       kb
1     2     1     0.11570           543920.0
2     3     1     0.14700           265265.6


[ angles ]
;  ai    aj    ak   funct        theta         kt
      1     2     3     1      180.0         1255.2
```

5- After solvent molecules are added into the simulation box, the *.top file is modified to include the number of solvent molecules in the system.

6- Define the *.ndx file. This file contains the atom index for each atom groups, for example, polymer atom group and solvent atom group.

7- The *.pdb, *.ndx, *.top, *.itp will be transferred to the computational computer (Palmetto) for the actual MD simulation. A bash_job file is also generated for this purpose. The content of the bash_job file is presented in the Section B.2.

The following bash script performs all seven steps mentioned above:

```
#!/bin/bash
```

```
echo 'Name of the .pdb file (‘temp’ in temp.pdb)?'
read TEMP

#run energy minimization in vacuum, solvate the molecules with
acetonitrile
#prepare the system, make the index file, and transfer to Palmetto

pdb2gmx -ff hopls -f ${TEMP}.pdb -p ${TEMP}.top -o ${TEMP}1.pdb -i
${TEMP}.itp
grompp -f sd.mdp -c ${TEMP}1.pdb -o ${TEMP}1.tpr  -p ${TEMP}.top
mdrun -v -deffnm ${TEMP}1
grompp -f cg.mdp -c ${TEMP}1.gro -o ${TEMP}2.tpr -p ${TEMP}.top
mdrun -v -deffnm ${TEMP}2

#This will open a graphic windows for visualization, when finished,
close the windows and the script would continue

ngmx -f ${TEMP}2.gro -s ${TEMP}2.tpr
editconf -f ${TEMP}2.gro -o ${TEMP}3.pdb -d 1.2 -c

#this script will add acetonitrile into the system, change the *.gro
and *.itp files to appropriate solvent.
genbox -cs acnua_4.gro -cp ${TEMP}3.pdb -o ${TEMP}_a.pdb -p ${TEMP}.top
2>&1 | tee test2 | egrep 'Added' > 1
sed -e 's/Added/AN1/' 1 | sed -e 's/molecules/ /' > 2
cp ${TEMP}.top 3
sed -e 's/spc.itp/acnua.itp/' 3 > 4
cat 4 2 > 5

# need ndx file

cat ndx | make_ndx -f ${TEMP}_a.pdb -o ${TEMP}_a.ndx
mv 5 ${TEMP}_a.top
sed -e "s/NAME/${TEMP}_a/g" bash_job > ${TEMP}_a
rm 1 2 3 4

#change the following address to the appropriate folder
scp ${TEMP}_a* ${TEMP}.itp
hhnguye@user.palmetto.clemson.edu:/common1/catalyst/hhnguye/
```

The content of the file *ndx* text file required to generate *.ndx in the above bash script:

```
1 | 2 | 3
del 1
del 1
del 1
name 2 PPE

q
```

B.2 Energy minimization of the simulated system

Before conducting MD or REMD simulations, the simulated system should be energy minimized to remove any high energy contacts and to achieve pseudo-equilibrium. The following bash file is used to conduct these steps.

The content of the bash_job text file required for the bash script presented in Section B.1:

```
#!/bin/bash -l
#PBS -l nodes=1:ppn=1
#PBS -q main
#PBS -l walltime=01:00:00
#PBS -j oe
#PBS -k n
#PBS -m n
#PBS -W x=\"QOS:General\"

module add intel/10.1
module add mpich/1.2.7
module add fftw
module add gromacs
export GMXLIB=/home/hhnguye/top
cd $PBS_O_WORKDIR

echo "$PBS_JOBNAME started at:"
date
echo "Use Gromacs-3.3 with double precision - Prior to MD"

#Energy minimization for the solvated box
  grompp_d -f sd.mdp -c NAME.pdb -p NAME.top -o NAME1.tpr
  mpirun -np 1 -machinefile $PBS_NODEFILE
/opt/gromacs/3.3.1/bin/mdrun_d -deffnm NAME1
# Energy minimization for solvated polymer molecule - L-bfgs algorithm
(em-tol= 10)
  grompp_d -f lbfgs.mdp -c NAME1.gro -p NAME.top -o NAME2.tpr
  mpirun -np 1 -machinefile $PBS_NODEFILE
/opt/gromacs/3.3.1/bin/mdrun_d -deffnm NAME2
# Position restrainst MD for 20ps to relax water molecules
  grompp_d -f prmdua.mdp -c NAME2.gro -p NAME.top -o NAME3.tpr -n
NAME.ndx
  mpirun -np 1 -machinefile $PBS_NODEFILE
/opt/gromacs/3.3.1/bin/mdrun_d -deffnm NAME3
# MD -NPT
  grompp_d -f ua.mdp -c NAME3.gro -p NAME.top -o NAME4.tpr -n NAME.ndx
```

```
  mpirun -np 1 -machinefile $PBS_NODEFILE
/opt/gromacs/3.3.1/bin/mdrun_d -deffnm NAME4
# MD - NVT
  grompp_d -f uaNVT.mdp -c NAME4.gro -p NAME.top -o NAME5.tpr -n
NAME.ndx
  mpirun -np 1 -machinefile $PBS_NODEFILE
/opt/gromacs/3.3.1/bin/mdrun_d -deffnm NAME5

echo "$PBS_JOBNAME finished at:"
date
```

For more information about the computational queue system or the modules available on Palmetto, see http://citi.clemson.edu/HPC.

B.3 The *.mdp file

The *.mdp file list the parameters used by Gromacs during a simulation. The most often used are the integration scheme (the intergrator), the time step (dt), the length of the simulation (nsteps), the neighbor update frequency (nlist), the trajectory saving frequency (nxout), the Coulombic interaction calculation method (coulomtype), the Lennard-Jones interaction calculation method (vdwtype), the energy atom groups (energygrps), the temperature coupling method (tcoupl), and the pressure coupling method (pcoupl). More information about these and other options can be found in the Gromacs documentation at www.gromacs.org/documentation.

The following is an example of a *.mdp file.

```
title       = PPE_md
cpp         = /lib/cpp
integrator  = md
tinit       = 0 ; ps
dt          = 0.002 ;ps
nsteps      = 5000000 ; total 10000ps=10ns
nstxout     = 5000
nstvout     = 0
nstfout     = 0
nstlog      = 5000
nstenergy   = 1000
nstxtcout   = 0
nstlist     = 10
ns_type     = grid
pbc         = xyz
rlist       = 1.2 ; nm
coulombtype = cut-off
rcoulomb    = 1.2 ;nm
vdwtype     = cut-off
rvdw        = 1.2 ; nm
energygrps  = PPE AN1
tcoupl      = berendsen
tc-grps     = PPE AN1
tau_t       = 0.1       0.1   ; ps
ref_t       = TEMP      TEMP  ; K
gen_vel     = yes
gen_temp    = TEMP ;K
gen_seed    = 173529
pcoupl          = Parrinello-Rahman
pcoupltype      = isotropic
tau_p           = 1.0
compressibility = 10.7e-5
ref_p           = 1
constraints = all-bonds
constraint_algorithm    = lincs
lincs_order = 4
lincs_warnangle   = 30
morse       = no
-----------------------------------------
```

B.4 Bash script for Replica Exchange Molecular Dynamics (REMD) simulation

For REMD simulations, it is required to define the temperature range and the temperature distribution. The temperarure distribution is calculated using the following equation (as proposed by former student Dr. Jay McAlilley):

$$T_i = T_{initial} \times \left(\frac{T_{final}}{T_{initial}} \right)^{i/max}, \quad i=0...max \qquad (B.1)$$

where T_i is the temperature of replica i, max is the total number of replicas in the REMD simulation, $T_{initial}$ and T_{final} are the two limits of the temperature range.

The following bash script is used to calculate the temperature distribution given the number of the replicas (1 CPU used for each replica) and the temperature range, then generate the *.tpr file for each replica at it's defined temperature, conduct an REMD run and calculate the time evolution of the radius of gyration of the simulated mPPE during simulation.

```
#!/bin/tcsh
#PBS -l nodes=1:ppn=1
#PBS -l walltime=00:01:00
#PBS -j oe
#PBS -M hhnguye@clemson.edu
#PBS -q main

module add intel/10.1
module add mpich/1.2.7
module add fftw
module add gromacs
setenv GMXLIB /home/hhnguye/top

cd $PBS_O_WORKDIR

set NCPUS=`grep -c ^ $PBS_NODEFILE`


echo "Running job on $NCPUS cpus"

set Ti=300
set Tf=600

@ maxrep = $NCPUS - 1
foreach rep (`seq 0 $maxrep`)
```

```
      set T=`echo "print
'%.2f'%($Ti.0*($Tf.0/$Ti.0)**($rep.0/$maxrep.0))"|python`
   sed -e "s/TEMP/$T/g" a.mdp > ${PBS_JOBNAME}$rep.mdp
   grompp_d -np 1 -f ${PBS_JOBNAME}$rep.mdp -c 24ENE_a5.gro -p *.top -n
*.ndx -o ${PBS_JOBNAME}$rep.tpr
end
mpirun -np $NCPUS -machinefile $PBS_NODEFILE
/opt/gromacs/3.3.1/bin/mdrun_d -deffnm ${PBS_JOBNAME}  -np $NCPUS -
multi -replex 250

rm *mdout.*
mkdir backup
mv ${PBS_JOBNAME}0.* ./backup
rm ${PBS_JOBNAME}*.*

cp *.ndx ./backup
cd backup

echo '2' | g_gyrate_d -f ${PBS_JOBNAME}0.trr -s ${PBS_JOBNAME}0.tpr -n
*.ndx -o gyrate_${PBS_JOBNAME}.xvg
echo " Finish $PBS_JOBNAME on"
date
```

B.5 Other post simulation analysis

1) Energy: use the *g_energy* command to calculate interaction between specific atom groups in the simulated system, such as electrostatic and Lennard-Jones interactions. Temperature and pressure are also calculated with *g_energy*.

2) Distance: the *g_distance* command calculates the time evolution of distance between the center of mass of two specific groups of atoms.

3) Bonds, angles, and dihedrals: use *g_bond*, *g_angle* and *g_dihedral* commands to monitor the time evolution of a specific bond, angle or dihedral angle during simulation.

4) Solvent accessible surface area: use *g_sas* to calculate the solvent asscessible surface area of the polymer.

5) Hydrogen bond: use *g_hbond* to conduct hydrogen bond analysis including the distance, angle, the number of bond between two specific atom groups.

6) Simulation trajectory visualization: use ngmx to open the trajectory files *.trr or *.xtc. Aletrnatively, other graphical programs like Visual Molecular Dynamic (VMD) can be use to visualize simulation trajectories. For further information about VMD, see http://www.ks.uiuc.edu/Research/vmd/.

APPENDIX C

MONITORING mPPE CONFORMATION TRANSITION

The parameters presented in Chapter 2 used in this study to monitor the transition of the mPPE conformation are not ordered structural parameters, and have to be supplemented by visual inspection to ensure correct observations. In this section, two new order or structural parameters are also introduced to easily quantify the extent to which helical mPPE structures have been formed. The first order parameter is the number of π-stacking aromatic rings and the second is the *cisoid* conformation number.

C.1 Python script to calculate the π-stacking aromatic rings

The number of consecutive π-stacking pairs can be used as an ordered parameter to monitor the formation of the helical secondary structure. In an mPPE helical conformation, the residues *i* and *i+6* are in overlapping positions when the approximate distance between rings (D_a) is about 0.45 nm. For example, Figure C.1 shows π-stacking pairs between the residues 1 and 7 as well as residues 2 and 8. From the number of consecutive π-stacking pairs, an mPPE conformation could be categorized as fully folded, half folded or random. A fully folded 12-mer mPPE will have 6 consecutive π-stacking pairs between residues *i* and *i+6*. The following python script is for this calculation.

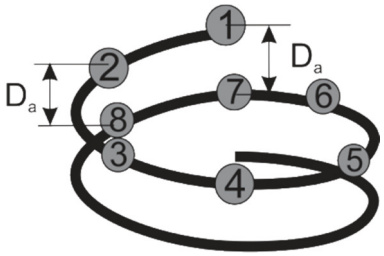


Figure C.1 Representative mPPE helical structure to illustrate presence of π-stacking pairs (between residues *i* and *i+6*). The distance between the residue of a π-stacking pair (D_a) is approximately 0.45 nm. In this figure π-stacking pair (1, 7) and (2, 8) are shown.

Python script to calculate the number of π-stacking pairs is listed below:

```
# Originally developed by Jay H. McAliley, Clemson Univeristy.
# Pymacs must be installed for this scripts to work properly.Pymacs
# version 0.1, http://wwwuser.gwdg.de/~dseelig/pymacs.html. This
# module need to be installed on Palmetto.
# The trr trajectory needs to be converted to an xtc trajectory
# use trjconv.
# !/usr/bin/python

import sys
from geometry import geometry
from pymacs import *
from math import sqrt
from optparse import OptionParser

# Parse Input Options
parser = OptionParser()
parser.add_option("-f", dest="xtc",
                        help=" Input trajectory file (.xtc)",
                        metavar=" FILE",
                        default=None)
parser.add_option("-a", dest="atoms",
                        help=" List of atoms in benzene rings",
                        metavar=" FILE",
                        default=None)
parser.add_option("-b", dest="begin",
```

```
                        help=" Time (ps) to begin calculating
averages",
                        type="float",
                        metavar=" TIME",
                        default=None)

(options, args) = parser.parse_args()
if None in [options.xtc,options.xvg,options.atoms]:
      parser.print_help()
      sys.exit()

# Set r2min
r2min=0.6**2      # nm^2

# Read atoms list
rings=open(options.atoms).readlines()
for i,ring in enumerate(rings):
      rings[i]=ring.split()
      for j,atm in enumerate(rings[i]):
            # XTC file object uses zero-based indices for atoms
            rings[i][j]=int(atm)-1
# Maximum fold number
maxfno=len(rings)-6

# Read .xtc file and calculate distances between 1,7 benzene rings
xtcfile=openXTC(options.xtc)
b40=''
for i in range(40):
      b40+='\b'
while True:
      # the coordinates are stored in frame['x']
      f = readXTCFrame(xtcfile)
      if not f:
            break
      frame=f

      # Are we past the specified beginning frame?
      sys.stderr.write('%40s%40s'%(b40,'reading frames, time
%.2f'%frame['time']))
      if frame['time']<options.begin:
            continue

      # Read box vectors
      B=[0,0,0]
      B[0]=frame['box'][0][0]
      B[1]=frame['box'][1][1]
      B[2]=frame['box'][2][2]

      # Calculate fold number
      fno=0
      coms=[]
      for ring in rings:
            x0=frame['x'][ring[0]]
            com=[0.0,0.0,0.0]
```

```
            for atm in ring:
                  x=frame['x'][atm]
                  for i in range(3):
                        # Remove periodicity
                        x[i]-=B[i]*round((x[i]-x0[i])/B[i])
                        com[i]+=x[i]
            # Append to center of mass list
            coms.append([])
            for i in range(3):
                  coms[-1].append(com[i]/6)
      # Calculate distances between 1,7 neighbors
      for i in range(6,len(coms)):
            r2=0.0
            for j in range(3):
                  # Remove periodicity
                  coms[i][j]-=B[j]*round((coms[i][j]-coms[i-
6][j])/B[j])
                  r2+=(coms[i][j]-coms[i-6][j])**2
            if r2<r2min:
                  # Increment fold number
                  fno+=1
      print frame['time'], fno
```

Python script to calculate the π-stacking aromatic rings, version 2.

```
#!/usr/bin/python
# Originally developed by Jay H. McAliley, Clemson Univeristy.
# Pymacs must be installed for this scripts to work properly. Pymacs
# version 0.1, http://wwwuser.gwdg.de/~dseelig/pymacs.html. This
# module need to be installed on Palmetto.
# The trr trajectory need to be converted to an xtc trajectory
# use trjconv.
# !/usr/bin/python

import sys
from geometry import geometry
from pymacs import *
from math import sqrt
from optparse import OptionParser

# Parse Input Options
parser = OptionParser()
parser.add_option("-f", dest="xtc",
                        help=" Input trajectory file (.xtc)",
                        metavar=" FILE",
                        default=None)
parser.add_option("-e", dest="xvg",
                        help=" Input energy file (.xvg)",
```

```
                        metavar=" FILE",
                        default=None)
parser.add_option("-a", dest="atoms",
                        help=" List of atoms in benzene rings",
                        metavar=" FILE",
                        default=None)
parser.add_option("-b", dest="begin",
                        help=" Time (ps) to begin calculating
averages",
                        type="float",
                        metavar=" TIME",
                        default=None)

(options, args) = parser.parse_args()
if None in [options.xtc,options.xvg,options.atoms]:
      parser.print_help()
      sys.exit()

# Set r2min
r2min=0.6**2      # nm^2
rhist={}
binw=0.1

# Read atoms list
rings=open(options.atoms).readlines()
for i,ring in enumerate(rings):
      rings[i]=ring.split()
      for j,atm in enumerate(rings[i]):
            # XTC file object uses zero-based indices for atoms
            rings[i][j]=int(atm)-1
# Maximum fold number
maxfno=len(rings)-6

# Read xvg file header and store series names
xvgfile=open(options.xvg)
series=[]
while True:
      xvgline=xvgfile.readline()
      if xvgline[0] not in ['@','#']:
            break
      elif 'legend "' in xvgline:
            series.append(xvgline.split()[-1][1:-1])

# Initialize statistics
sx,ssx,N=([],[],[])
for fno in range(maxfno+1):
      sx.append([])
      ssx.append([])
      N.append(0)
      for i in range(len(series)):
            sx[-1].append(0.0)
            ssx[-1].append(0.0)

# Read .xtc file and calculate distances between 1,7 benzene rings
```

```
xtcfile=openXTC(options.xtc)
b40=''
for i in range(40):
      b40+='\b'
while True:
      # the coordinates are stored in frame['x']
      f = readXTCFrame(xtcfile)
      if not f:
            break
      frame=f

      # Are we past the specified beginning frame?
      sys.stderr.write('%40s%40s'%(b40,'reading frames, time
%.2f'%frame['time']))
      if frame['time']<options.begin:
            continue

      # Read box vectors
      B=[0,0,0]
      B[0]=frame['box'][0][0]
      B[1]=frame['box'][1][1]
      B[2]=frame['box'][2][2]

      # Calculate fold number
      fno=0
      coms=[]
      for ring in rings:
            x0=frame['x'][ring[0]]
            com=[0.0,0.0,0.0]
            for atm in ring:
                  x=frame['x'][atm]
                  for i in range(3):
                        # Remove periodicity
                        x[i]-=B[i]*round((x[i]-x0[i])/B[i])
                        com[i]+=x[i]
            # Append to center of mass list
            coms.append([])
            for i in range(3):
                  coms[-1].append(com[i]/6)
      # Calculate distances between 1,7 neighbors
      for i in range(6,len(coms)):
            r2=0.0
            for j in range(3):
                  # Remove periodicity
                  coms[i][j]-=B[j]*round((coms[i][j]-coms[i-
6][j])/B[j])
                  r2+=(coms[i][j]-coms[i-6][j])**2
            if r2<r2min:
                  # Increment fold number
                  fno+=1
            r=sqrt(r2)
            bin=int(r//binw)
            if bin in rhist:
                  rhist[bin]+=1
```

```
            else:
                  rhist[bin]=1

      # Get energy values
      while int(1e6*float(xvgline.split()[0])) !=
int(1e6*frame['time']):
            xvgline=xvgfile.readline()
      for i,val in enumerate(xvgline.split()[1:]):
            val=float(val)
            sx[fno][i]+=val
            ssx[fno][i]+=val**2
      N[fno]+=1

# Print results
sys.stderr.write('\n')
line='%20s%20s'%('Fold Number','N')
for i,s in enumerate(series):
      line+='%20s%20s'%(s,'st.dev.')
print line
avgfno=0.0
Ntot=0
for fno in range(maxfno+1):
      line='%20i%20i'%(fno,N[fno])
      avgfno+=fno*N[fno]
      Ntot+=N[fno]
      for i,s in enumerate(series):
            if N[fno]!=0:
                  avg=sx[fno][i]/N[fno]
                  stdev=sqrt(ssx[fno][i]/N[fno]-avg**2)
                  line+='%20.6f%20.6f'%(avg,stdev)
            else:
                  line+='%20s%20s'%('','')
      print line
if Ntot!=0:
      avgfno/=Ntot
print '\nAverage Fold Number:%20.2f'%(avgfno)

print '\nHistogram for r-values:\n%20s%20s'%('r (nm)','freq')
keys=rhist.keys()
keys.sort()
for i in range(keys[-1]):
      if i in rhist:
            val=rhist[i]
      else:
            val=0.0
      print '%20.6f%20i'%(binw*i,val)
```

Bash file to run the above python scripts

```
#!/bin/tcsh
#PBS -l nodes=1:ppn=1
#PBS -l walltime=01:00:00
#PBS -j oe
#PBS -M raymons@clemson.edu
#PBS -q main

module add intel/10.1
module add mpich/1.2.7
module add fftw
module add gromacs/4.0.5
setenv GMXLIB /home/hhnguye/top
setenv PYTHONPATH /projsmall/catalyst/share


#1. Convert the trajectories file .trr to .xtc
#2. Prepare the index file contain only aromatic carbon number of the
polymer chain
#each line contain only 6 carbon of 1 aromatic ring
#the number of lines equals to the number of mers of the polymer chain
#see example aromatics_C.ndx, this file is prepared for 12 mer mPPE, it
has 12 lines for 12 aromatic rings, each line has 6 id numbers for
#6 aromatic carbon (in mPPE model, aromatic carbons are named C7, C8,
C10, C11, C13, C15)
#3.Replace the FILE_NAME in the python command line
#3.Output file has      "Time" and "Fold number"
#to be considered as folded, the Fold number needs to be equal or
greater than half of the mer number

cd $PBS_O_WORKDIR
echo 'start at:'
date

pythonn  foldNumber_v1.py -f FILE_NAME.xtc -a FILE_NAME.ndx -b
${BEGIN_TIME} > foldnumber.xvg

end
date
```

C.2 The *cisoid* number

The number of consecutive *cisoid* arrangements of aromatic rings can be used as an ordered parameter to monitor the formation of helical mPPE structures. In an mPPE helical conformation, the residues *i* and *i+3* are in *cisoid* conformation, same side of the

line connecting residues *i+1* and *i+2*. For example, in Figure C.2, with the four consecutive residues, 1, 2, 3 and 4, in a helical structure, the residues 1 and 4 are in *cisoid* conformation; whereas in a random extended structure, the residues 1 and 4 (or other non-*cisoid*) are in *transoid* conformation.

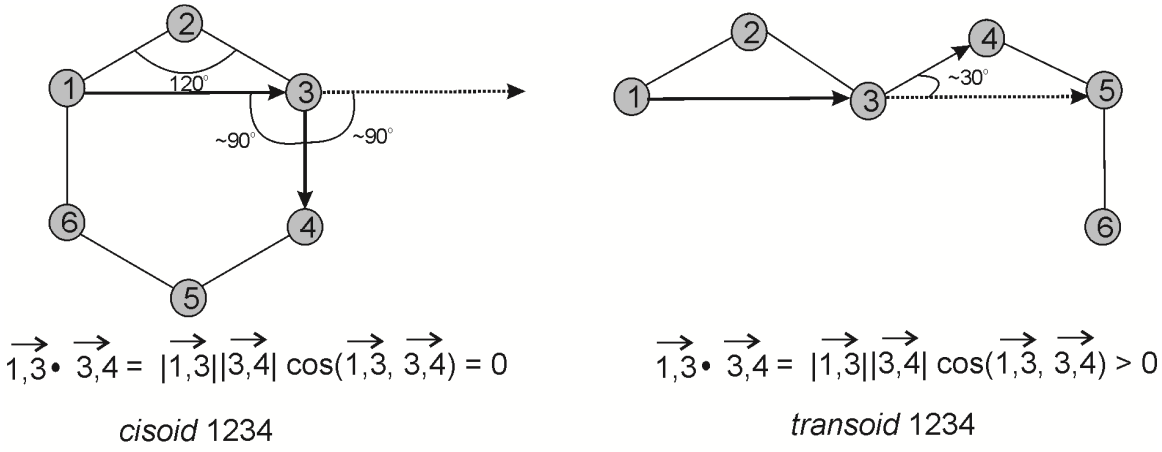


Figure C.2 Illustration of the c*isoid* number calculation scheme. In the *cisoid* conformation, the residues 1 and 4 are located on the same side with respect to the line formed between residues 2 and 3, whereas in the *transoid* conformation, the residues 1 and 4 are located at the opposite sides in respect to the line formed between residues 2 and 3. The dot procducts of vectors $\vec{1,3}$ and $\vec{3,4}$ in *cisoid* and *transiod* conformations are also shown. The dashed vectors are to show the $\vec{1,3}$ vector shifted to have the same initial point as vector $\vec{3,4}$.

The conformation of residues *i* and *i+3* is calculated as by determining the dot product between the vector $\vec{i, i+2}$, connecting the centers of mass of residues *i* and *i+2*, and the vector $\vec{i+2, i+3}$, connecting the centers of mass of residues *i+2* and *i+3*. If this dot

product is equal to zero, residues *i* and *i+3* are in *cisoid* conformation, whereas this dot product is greater than zero, residues *i* and *i+3* are in a *non-cisoid* or *transoid* conformation. For example, in Figure C.2, if residues 1 and 4 are in *cisoid* conformation, the dot product of the vector connecting bthe centers of mass of residues 1 and 3, ($\overrightarrow{1,3}$) with the vector connecting the centers of mass of residues 3 and 4, ($\overrightarrow{3,4}$) equals to zero. On the other hand, if the conformation is *transoid*, the dot product of the vector connecting between the centers of mass of residue 1 and 3, or $\overrightarrow{1,3}$, with the vector connecting between the centers of mass of residue 3 and 4, or $\overrightarrow{3,4}$, is greater than zero.

The Cartesian coordinates of the vectors $\overrightarrow{i, i+2}$ and $\overrightarrow{i+2, i+3}$ can be calculated from the following equations:

$$\overrightarrow{i, i+2} = | (x_i - x_{i+2}) \ (y_i - y_{i+2}) \ (z_i - z_{i+2}) | \qquad \text{(D.1)}$$

$$\overrightarrow{i+2, i+3} = | (x_{i+2} - x_{i+3}) \ (y_{i+2} - y_{i+3}) \ (z_{i+2} - z_{i+3}) | \qquad \text{(D.2)}$$

where (x_i, y_i, z_i), $(x_{i+2}, y_{i+2}, z_{i+2})$ and $(x_{i+3}, y_{i+3}, z_{i+3})$ are the Cartesian coordinates of the centers of mass of residues i, i+2 and i+3, respectively.

The dot product between vectors $\overrightarrow{i, i+2}$ and $\overrightarrow{i+2, i+3}$ can be calculated using the following equation:

$$\overrightarrow{i, i+2} \bullet \overrightarrow{i+2, i+3} = x_{\overrightarrow{i,i+2}} x_{\overrightarrow{i+2,i+3}} + y_{\overrightarrow{i,i+2}} y_{\overrightarrow{i+2,i+3}} + z_{\overrightarrow{i,i+2}} z_{\overrightarrow{i+2,i+3}} \qquad \text{(D.3)}$$

Similar to the calculation of the number of consecutive π-stacking pairs, determining the number of consecutive *cisoid* conformations enables one to categorize an mPPE conformation as fully folded, half folded or random. One complete helix turn would have 3 consecutive *cisoid* conformations between residues *i* and *i+3*.

C.3 Method to determine the direction of the helix turn: the Q number

During a given simulation an mPPE can fold into eith of two types of helical conformations, counter clock wise helices or clockwise helices (as shown in Figure D.3). This section presents a method to distinguih between these two enantiomers by calculating the cross product and dot product of the three vectors $\overrightarrow{i, i+1}$, $\overrightarrow{i+1, i+2}$ and $\overrightarrow{i, i+6}$ via the following equation:

$$Q = (\overrightarrow{i, i+1} \times \overrightarrow{i+1, i+2}) \bullet \overrightarrow{i, i+6} \qquad \text{(D.4)}$$

where $\overrightarrow{i, i+1}$ is the vector connecting the centers of mass between residues *i*, *i+1*, $\overrightarrow{i+1, i+2}$ is the vector connecting the centers of mass between residues *i+1*, *i+2*, and $\overrightarrow{i, i+6}$ is the vector connecting the centers of mass between residues *i*, *i+6*.

The Cartesian coordinates of the vectors $\overrightarrow{i, i+1}$, $\overrightarrow{i+1, i+2}$ and $\overrightarrow{i, i+6}$ can be calculated from the following equations:

$$\overrightarrow{i, i+1} = | (x_i - x_{i+1}) \ (y_i - y_{i+1}) \ (z_i - z_{i+1}) | \qquad \text{(D.5)}$$

$$\overrightarrow{i+1,\, i+2} = |\,(x_{i+1}-x_{i+2})\;\;(y_{i+1}-y_{i+2})\;\;(z_{i+1}-z_{i+2})\,| \qquad \text{(D.6)}$$

$$\overrightarrow{i,\, i+6} = |\,(x_{i}-x_{i+6})\;\;(y_{i}-y_{i+6})\;\;(z_{i}-z_{i+6})\,| \qquad \text{(D.7)}$$

where (x_i, y_i, z_i), $(x_{i+1}, y_{i+1}, z_{i+1})$, $(x_{i+2}, y_{i+2}, z_{i+2})$ and $(x_{i+6}, y_{i+6}, z_{i+6})$ are the Cartesian coordinates of the centers of mass of residues *i*, *i+1*, *i+2* and *i+6*, respectively.

Using Eqn. (D.4) to (D.7), the quantity Q can be calculated using the following equation:

$$Q = (\overrightarrow{i,\, i+1} \times \overrightarrow{i+1,\, i+2}) \bullet \overrightarrow{i,\, i+6} = \begin{vmatrix} x_{\overline{i,i+1}} & y_{\overline{i,i+1}} & z_{\overline{i,i+1}} \\ x_{\overline{i+1,i+2}} & y_{\overline{i+1,i+2}} & z_{\overline{i+1,i+2}} \\ x_{\overline{i,i+6}} & y_{\overline{i,i+6}} & z_{\overline{i,i+6}} \end{vmatrix} \qquad \text{(D.8)}$$

The quantities Q is less than zero if the helix rotates in a counter-clock-wise direction and Q is greater than zero if the helix rotates in a clockwise direction.

For example, in Figure D.3, on the left, the helix turns in a counter clock wise direction:

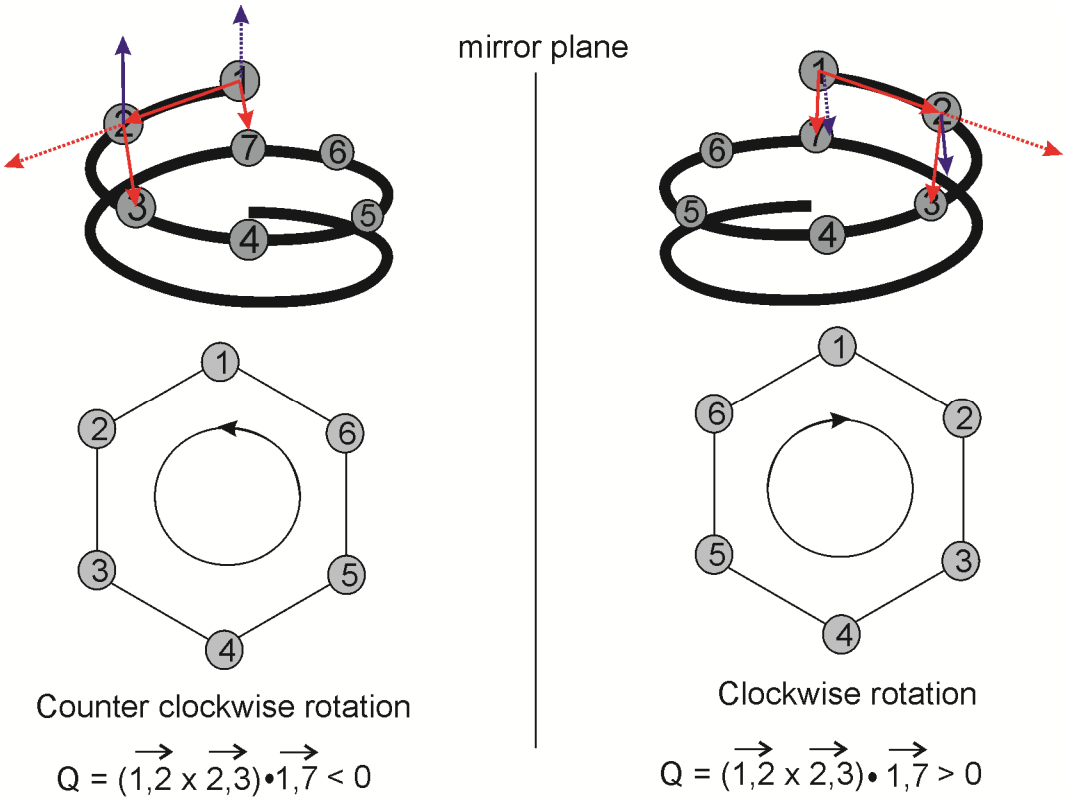


Figure C.3 Scheme for determining the chirality of an mPPE helix. By calculating the cross product and dot product of vectors $\overrightarrow{1,2}$, $\overrightarrow{2,3}$ and $\overrightarrow{1,7}$ (red solid vectors), the direction of the helix turn can be determined. For the counter clockwise rotating helix, $Q = (\overrightarrow{1,2} \times \overrightarrow{2,3}) \bullet \overrightarrow{1,7} < 0$, whereas for the clockwise rotating helix, $Q = (\overrightarrow{1,2} \times \overrightarrow{2,3}) \bullet \overrightarrow{1,7} > 0$. The cross product vectors $(\overrightarrow{1,2} \times \overrightarrow{2,3})$ are shown as purple color vectors. The dashed red vector is the vector $\overrightarrow{1,2}$ shifted to have the same initial point as vector $\overrightarrow{2,3}$. The dashed purple vector is the cross product $(\overrightarrow{1,2} \times \overrightarrow{2,3})$ shifted to have the same initial point as vector $\overrightarrow{1,7}$.

$$Q = (\overrightarrow{1, 2} \times \overrightarrow{2, 3}) \bullet \overrightarrow{1, 7} < 0 \qquad \text{(D.9)}$$

On the right, the helix turns in a clock wise direction:

$$Q = (\overrightarrow{1, 2} \times \overrightarrow{2, 3}) \bullet \overrightarrow{1, 7} > 0 \qquad \text{(D.10)}$$

This calculation should be best coupled with the other monitoring parameters, including the consecutive π-stacking number or the consecutive *cisoid* number. During a simulation, the consecutive π-stacking number and the consecutive *cisoid* number will help to determine if the polymer is in a helical conformation, while the Q number will determine the optical activity of the helix (i.e.,. which enantiomer of the helix formed).

APPENDIX D

ESTER FUNCTIONALIZED mPPE FORCE FIELD

D.1 OPLS force field

A force field is a set of parameters from which a simulation program calculates the forces acting on each particle during a molecular dynamics simulation. In Gromacs, the force field parameters are listed in nine files whose names start with two letters *ff*. For example, the OPLS force field consists of nine files named *ffopls*.*. The function of each file in the OPLS force field is described below:

1- ff*.itp: bonded and non-bonded parameter file.
2- ff*.atp: file containing atom types. For example, CA is the atom type of a carbon atom in an aromatic compound.
3- ff*.bon.itp: file containing bonded parameters: bonds (length, force constants), angles (angle, force constants), dihedral (dihedral angle, constants), improper diherals (improper dihedral angles, force constants).
4- ff*nb.itp: file containing non-bonded parameters for each atom type listed in ff*.atp: atom partial charge, Lennard-Jones constants (σ and ε).
5- ff*.hdb: file containing the hydrogen atom database including H connectivity, the number of hydrogen atoms connected, and the type of hydrogen atoms connected. If this hydrogen database is used to set up a simulation system, Gromacs will

ignore all specified hydrogen atoms and generate new hydrogen atoms for the imported structure. This hydrogen database is handy in building systems having large numbers of hydrogen atoms. By defining this hydrogen database, only the position of the heavy atoms (carbon, oxygen, nitrogen, …) need to be specified.

6- ff*.rtp: file containing structural information about the simulated molecules including the number of atoms, type of atoms used, how they arrange and connect to each other (bonds, angle, dihedral, improper). An example of this structural information (an ff*.rtp) file is presented in Section A.2.
7- ff*-c.tdb: file containing carbon termination. It is only needed for protein simulations.
8- ff*-n.tdb: file containing nitrogen information. It is only needed for protein simulations.
9- ff*.ddb: file containing dummy atoms. A dummy atom is a type of imaginary atom that is used to add constraint or forces on the simulated molecules.

More information can be found in the online Gromacs documentation located at www.gromacs.org/documentation. The following section describes the structural information for the ester functionalized mPPE model.

D.2 The ester functionalized mPPE model

The structural information of the ester functionalized mPPE model is listed in ff*.rtp files. The ester functionalized mPPE requires three sub-structures models, for the head,

body, and tail sections of the polymers. The three sub-structure models are also used for other functionalized mPPE models. In Gromacs, each model is called a residue. The atom names, atom arrangement, and connectivity for each sub-structural model are shown in Figure D.1.

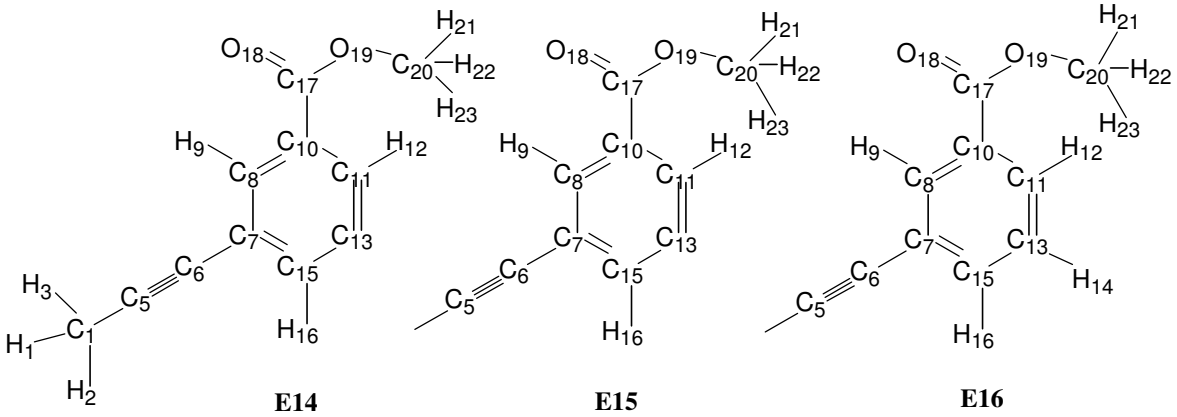


Figure D.1 Ester functionalized mPPE models with atom labels. Models E14, E15 and E16 are for the head, body, and tail sections of the ester functionalized mPPE, respectively. The atom C13 in the residues E14 and E15; as well as the atom C5 in residues E15 and E16 are at the connecting positions on the mPPE back bone.

The parameters for the ester functionalized mPPE model are listed in five sections in ff*.rtp files:

1- Name of the residue. For example, the name of residue E14

```
; Ester functionalized mPPE models
      [ E14 ]
```

2- atoms: list the name, type, partial charge and charge group of each atom. For example, information for atoms H1 and H2 of reside E14

```
[ atoms ]
; name type       charge   chargegroup
   H1  opls_140   0.060     1
   H2  opls_140   0.060     1
```

3- bonds: list all bonds and the bond force constant. For example, bond between H1 and C4

```
[ bonds ]
ai    aj        b0           kb
C4    H1     0.10900     284702.4
```

4- angles: list all angles. For example, angle H1, C4, C5

```
[ angles ]

;  ai   aj    ak        th0          cth
   H1   C4    C5      108.500      293.076
```

5- dihedrals: list all dihedrals. For example, dihedral H1, C4, C5, C6

```
[ dihedrals ]
;  ai   aj     ak    al          coefficients
   H1    C4       C5     C6    0.0  0.0  0.0  0.0  0.0  0.0
```

6- improper dihedrals: list all improper dihedrals. For example, C6, C8, C15, C7 (aromatic ring)

```
[ impropers ]
; ai   aj     ak    al       th0         cth
  C6    C8    C15    C7    0           334.72
```

D.3 The file ff*.rtp

```
; Ester functionalized mPPE models
 [ E14 ]
 [ atoms ]
; name              type    charge          chargegroup
   H1  opls_140             0.060               1
   H2  opls_140             0.060               1
```

```
   H3  opls_140            0.060              1
   C4  opls_135           -0.180              1
   C5  opls_931            0.000              1
   C6  opls_931            0.000              1
   C7  opls_145            0.000              2
   C8  opls_145           -0.115              2
   H9  opls_146            0.115              2
   C11 opls_145           -0.115              2
   H12 opls_146            0.115              2
   C13 opls_145            0.000              2
   C15 opls_145           -0.115              2
   H16 opls_146            0.115              2
   C10 opls_145           -0.115              3
   C17 opls_471            0.625              3
   O18 opls_466           -0.430              3
   O19 opls_467           -0.330              3
   C20 opls_468            0.160              3
   H21 opls_469            0.030              3
   H22 opls_469            0.030              3
   H23 opls_469            0.030              3
[ bonds ]
;  ai    aj        b0           kb
   C4    H1     0.10900     284702.4
   C4    H2     0.10900     284702.4
   C4    H3     0.10900     284702.4
   C5    C4     0.14700     326570.4
   C6    C5     0.12100     962964.0
   C7    C8     0.14000     392721.8
   C7    C6     0.14510     334944.0
   C8    C10    0.14000     392721.8
   C8    H9     0.10800     307311.1
   C10   C11    0.14000     392721.8
   C11   C13    0.14000     392721.8
   C11   H12    0.10800     307311.1
   C13   C15    0.14000     392721.8
   C15   C7     0.14000     392721.8
   C15   H16    0.10800     307311.1
   C10   C17    0.14900     334720.0
   C17   O18    0.12290     476976.0
   C17   O19    0.13270     179075.2
   O19   C20    0.14100     267776.0
   C20   H21    0.10900     284512.0
   C20   H22    0.10900     284512.0
   C20   H23    0.10900     284512.0
   C13   +C5    0.14510     334944.0
[ angles ]
;  ai   aj    ak        th0          cth
   H1   C4    C5      108.500      293.076
   H2   C4    C5      108.500      293.076
   H2   C4    H1      107.800      276.329
   H3   C4    C5      108.500      293.076
   H3   C4    H1      107.800      276.329
   H3   C4    H2      107.800      276.329
   C4   C5    C6      180.000     1256.040
```

```
   C5   C6    C7      180.000     1339.776
   C8   C7    C15     120.000      527.537
   C6   C7    C15     120.000      586.152
   C6   C7    C8      120.000      586.152
   C10  C8    C7      120.000      527.537
   H9   C8    C7      120.000      293.076
   H9   C8    C10     120.000      293.076
   C11  C10   C8      120.000      527.537
   C13  C11   C10     120.000      527.537
   H12  C11   C13     120.000      293.076
   H12  C11   C10     120.000      293.076
   C13  C15   C7      120.000      527.537
   H16  C15   C7      120.000      293.076
   H16  C15   C13     120.000      293.076
   C11  C13   C15     120.000      527.537
   C8   C10   C17     120.000      711.280
   C11  C10   C17     120.000      711.280
   C10  C17   O18     120.400      669.440
   C10  C17   O19     111.400      677.808
   O18  C17   O19     123.400      694.544
   C17  O19   C20     116.900      694.544
   O19  C20   H21     109.500      292.880
   O19  C20   H22     109.500      292.880
   O19  C20   H23     109.500      292.880
   H21  C20   H22     107.800      276.144
   H21  C20   H23     107.800      276.144
   H22  C20   H23     107.800      276.144
   C11  C13  +C5      120.000      586.152
   C15  C13  +C5      120.000      586.152
   C13 +C5   +C6      180.000     1339.776
[ dihedrals ]
;  ai   aj     ak    al          coefficients
   H1    C4       C5     C6    0.0  0.0  0.0  0.0  0.0  0.0
   H2    C4       C5     C6    0.0  0.0  0.0  0.0  0.0  0.0
   H3    C4       C5     C6    0.0  0.0  0.0  0.0  0.0  0.0
   C4    C5       C6     C7    0.0  0.0  0.0  0.0  0.0  0.0
   C10   C11      C13   +C5   30.334  0  -30.334 0  0  0
   C7    C8       C9    H10   30.334  0  -30.334 0  0  0
   H10   C9       C10   C17   30.334  0  -30.334 0  0  0
   C17   C10      C11   H12   30.334  0  -30.334 0  0  0
   H12   C11      C13   +C5   30.334  0  -30.334 0  0  0
   +C5   C13      C15   H16   30.334  0  -30.334 0  0  0
   H16   C15      C8     C7   30.334  0  -30.334 0  0  0
   C13   +C5      +C6   +C7    0.0  0.0  0.0  0.0  0.0  0.0
   C8    C10      C17   O18   8.78640     0.0   -8.78640  0.0 0.0 0.0
   C8    C10      C17   O19   8.78640     0.0   -8.78640  0.0  0.0 0.0
   C11   C10      C17   O18   8.78640     0.0   -8.78640  0.0  0.0 0.0
   C11   C10      C17   O19   8.78640     0.0   -8.78640  0.0  0.0 0.0
   C10   C17      O19   C20     29.288  -8.368 -20.92   0.0   0.0   0.0
   O18   C17      O19   C20   21.43881 0.0 -21.43881 0.0 0.0 0.0
   C17   O19      C20   H21   0.41421 1.24265 0.0 -1.65685 0.0 0.0
   C17   O19      C20   H22   0.41421 1.24265 0.0 -1.65685 0.0 0.0
   C17   O19      C20   H23   0.41421 1.24265 0.0 -1.65685 0.0 0.0000
```

```
   ; i      j     k     l     (ethynylene dihedral angles, k = 1196.2
J/mol, ~0.6kcal.mol)
  C11   C13 +C7   +C8   180   1196.2            2
[ impropers ]
; ai   aj     ak    al       th0         cth
   C6    C8    C15    C7    0           334.72
   C15   C7    C8    C10    0           334.72
   C7    C8    C10   C11    0           334.72
   C8    C10   C11   C13    0           334.72
   C8    C7    C15   C13    0           334.72
   C10   C11   C13   C15    0           334.72
   C7    C15   C13   C11    0           334.72
   C13   C11   C15   +C5    0           334.72
   H9    C10   C7    C8     0           334.72
   H12   C10   C13   C11    0           334.72
   H16   C13   C7    C15    0           334.72
   C17   C11   C8    C10    0           334.72

[ E15 ]
 [ atoms ]
; name              type    charge          chargegroup
   C5  opls_931             0.000              1
   C6  opls_931             0.000              1
   C7  opls_145             0.000              2
   C8  opls_145            -0.115              2
   H9  opls_146             0.115              2
   C11 opls_145            -0.115              2
   H12 opls_146             0.115              2
   C13 opls_145             0.000              2
   C15 opls_145            -0.115              2
   H16 opls_146             0.115              2
   C10 opls_145            -0.115              3
   O18 opls_466            -0.430              3
   O19 opls_467            -0.330              3
   C20 opls_468             0.160              3
   H21 opls_469             0.030              3
   H22 opls_469             0.030              3
   H23 opls_469             0.030              3
[ bonds ]
;  ai    aj        b0           kb
   C5   -C13    0.14700     326570.4
   C6    C5     0.12100     962964.0
   C7    C8     0.14000     392721.8
   C7    C6     0.14510     334944.0
   C8    C10    0.14000     392721.8
   C8    H9     0.10800     307311.1
   C10   C11    0.14000     392721.8
   C11   C13    0.14000     392721.8
   C11   H12    0.10800     307311.1
   C13   C15    0.14000     392721.8
   C15   C7     0.14000     392721.8
   C15   H16    0.10800     307311.1
   C10   C17    0.14900     334720.0
   C17   O18    0.12290     476976.0
```

```
   C17   O19    0.13270     179075.2
   O19   C20    0.14100     267776.0
   C20   H21    0.10900     284512.0
   C20   H22    0.10900     284512.0
   C20   H23    0.10900     284512.0
   C13   +C5    0.14510     334944.0
[ angles ]
;  ai   aj    ak        th0          cth
  -C13  C5    C6      180.000      1256.040
  -C11 -C13   C5      120.000      586.152
  -C15 -C13   C5      120.000      586.152
   C5   C6    C7      180.000      1339.776
   C8   C7    C15     120.000      527.537
   C6   C7    C15     120.000      586.152
   C6   C7    C8      120.000      586.152
   C10  C8    C7      120.000      527.537
   H9   C8    C7      120.000      293.076
   H9   C8    C10     120.000      293.076
   C11  C10   C8      120.000      527.537
   C13  C11   C10     120.000      527.537
   C13  C15   C7      120.000      527.537
   H12  C11   C13     120.000      293.076
   H12  C11   C10     120.000      293.076
   H16  C15   C7      120.000      293.076
   H16  C15   C13     120.000      293.076
   C11  C13   C15     120.000      527.537
   C8   C10   C17     120.000      711.280
   C11  C10   C17     120.000      711.280
   C10  C17   O18     120.400      669.440
   C10  C17   O19     111.400      677.808
   O18  C17   O19     123.400      694.544
   C17  O19   C20     116.900      694.544
   O19  C20   H21     109.500      292.880
   O19  C20   H22     109.500      292.880
   O19  C20   H23     109.500      292.880
   H21  C20   H22     107.800      276.144
   H21  C20   H23     107.800      276.144
   H22  C20   H23     107.800      276.144
   C11  C13   +C5     120.000      586.152
   C15  C13   +C5     120.000      586.152
   C13  +C5   +C6     180.000     1339.776

[ dihedrals ]
;  ai   aj     ak    al          coefficients
  -C13   C5       C6     C7    0  0  0 0  0  0
   C10   C11      C13   +C5   30.334  0  -30.334 0  0  0
   C7    C8       C9    H10   30.334  0  -30.334 0  0  0
   H10   C9       C10   C17   30.334  0  -30.334 0  0  0
   C17   C10      C11   H12   30.334  0  -30.334 0  0  0
   H12   C11      C13   +C5   30.334  0  -30.334 0  0  0
   +C5   C13      C15   H16   30.334  0  -30.334 0  0  0
   H16   C15      C8     C7   30.334  0  -30.334 0  0  0
   C13   +C5      +C6   +C7    0       0   0      0  0  0
   C8    C10      C17   O18    8.78640 0   0 -8.78640  0  0 0
```

```
   C8    C10      C17   O19    8.78640    0     -8.78640  0  0 0
   C11   C10      C17   O18    8.78640    0     -8.78640  0  0 0
   C11   C10      C17   O19    8.78640    0     -8.78640  0  0 0
   C10   C17      O19   C20   29.2880  -8.3680 -20.920  0   0   0
   O18   C17      O19   C20   21.43881 0.0 -21.43881 0.0 0.0 0.0
   C17   O19      C20   H21   0.41421 1.24265 0.0 -1.65685 0.0 0.0
   C17   O19      C20   H22   0.41421 1.24265 0.0 -1.65685 0.0 0.0
   C17   O19      C20   H23   0.41421 1.24265 0.0 -1.65685 0.0 0.0

;i    j     k     l     (ethynylene dihedral angles, k = 1196.2 J/mol,
~0.6kcal.mol)
   C11   C13      +C7   +C8         180   1196.2            2

[ impropers ]
; ai   aj     ak    al       th0         cth
   C6    C8    C15    C7    0           334.72
   C15   C7    C8    C10    0           334.72
   C7    C8    C10   C11    0           334.72
   C8    C10   C11   C13    0           334.72
   C8    C7    C15   C13    0           334.72
   C10   C11   C13   C15    0           334.72
   C7    C15   C13   C11    0           334.72
   C13   C11   C15  +C5     0           334.72
   H9    C7    C10   C8     0           334.72
   H12   C13   C10   C11    0           334.72
   H16   C13   C7    C15    0           334.72
   C17   C11   C8    C10    0           334.72
   C5   -C15  -C11  -C13    0           334.72

[ E16]
[ atoms ]
; name              type    charge          chargegroup
   C5  opls_931             0.000              1
   C6  opls_931             0.000              1
   C7  opls_145             0.000              2
   C8  opls_145            -0.115              2
   H9  opls_146             0.115              2
   C11 opls_145            -0.115              2
   H12 opls_146             0.115              2
   C13 opls_145            -0.115              2
   H14 OPLS_146             0.115              2
   C15 opls_145            -0.115              2
   H16 opls_146             0.115              2
   C10 opls_145            -0.115              3
   C17 opls_471             0.625              3
   O18 opls_466            -0.430              3
   O19 opls_467            -0.330              3
   C20 opls_468             0.160              3
   H21 opls_469             0.030              3
   H22 opls_469             0.030              3
   H23 opls_469             0.030              3

[ bonds ]
;  ai    aj        b0           kb
```

```
   C5   -C13    0.14700     326570.4
   C6    C5     0.12100     962964.0
   C7    C8     0.14000     392721.8
   C7    C6     0.14510     334944.0
   C8    C10    0.14000     392721.8
   C8    H9     0.10800     307311.1
   C10   C11    0.14000     392721.8
   C11   C13    0.14000     392721.8
   C11   H12    0.10800     307311.1
   C13   C15    0.14000     392721.8
   C15   C7     0.14000     392721.8
   C15   H16    0.10800     307311.1
   C10   C17    0.14900     334720.0
   C17   O18    0.12290     476976.0
   C17   O19    0.13270     179075.2
   O19   C20    0.14100     267776.0
   C20   H21    0.10900     284512.0
   C20   H22    0.10900     284512.0
   C20   H23    0.10900     284512.0
   C13   H14    0.10800     307311.1 ; END OF THE POLYMER

[ angles ]
;  ai   aj    ak        th0          cth
  -C13  C5    C6      180.000      1256.040
  -C11 -C13   C5      120.000      586.152
  -C15 -C13   C5      120.000      586.152
   C5   C6    C7      180.000      1339.776
   C8   C7    C15     120.000      527.537
   C6   C7    C15     120.000      586.152
   C6   C7    C8      120.000      586.152
   C10  C8    C7      120.000      527.537
   H9   C8    C7      120.000      293.076
   H9   C8    C10     120.000      293.076
   C11  C10   C8      120.000      527.537
   C13  C11   C10     120.000      527.537
   C13  C15   C7      120.000      527.537
   H12  C11   C13     120.000      293.076
   H12  C11   C10     120.000      293.076
   H16  C15   C7      120.000      293.076
   H16  C15   C13     120.000      293.076
   C11  C13   C15     120.000      527.537
   C8   C10   C17     120.000      711.280
   C11  C10   C17     120.000      711.280
   C10  C17   O18     120.400      669.440
   C10  C17   O19     111.400      677.808
   O18  C17   O19     123.400      694.544
   C17  O19   C20     116.900      694.544
   O19  C20   H21     109.500      292.880
   O19  C20   H22     109.500      292.880
   O19  C20   H23     109.500      292.880
   H21  C20   H22     107.800      276.144
   H21  C20   H23     107.800      276.144
   H22  C20   H23     107.800      276.144
   C11  C13   H14     120.000      293.076
```

```
   C15  C13   H14     120.000      293.076


[ dihedrals ]
;  ai   aj     ak    al          coefficients
  -C13   C5       C6    C7    0.0  0.0  0.0 0.0  0.0  0.0
   C7    C8       C9    H10   30.334  0.0  -30.334 0  0  0
   H10   C9       C10   C17   30.334  0.0  -30.334 0  0  0
   C17   C10      C11   H12   30.334  0.0  -30.334 0  0  0
   H16   C15      C8    C7    30.334  0.0  -30.334 0  0  0
   C8    C10      C17   O18   8.78640 0.0 -8.78640  0.0  0.0 0.0
   C8    C10     C17    O19   8.78640 0.0 -8.78640  0.0  0.0 0.0
   C11   C10     C17    O18   8.78640 0.0 -8.78640  0.0  0.0 0.0
   C11   C10      C17   O19   8.78640 0.0 -8.78640  0.0  0.0 0.0
   C10   C17      O19   C20   29.2880 -8.3680 -20.920   0.0   0.0   0.0
   O18   C17      O19   C20   21.43881 0.0 -21.43881 0.0 0.0 0.0
   C17   O19      C20   H21   0.41421 1.24265 0.0 -1.65685 0.0 0.0
   C17   O19      C20   H22   0.41421 1.24265 0.0 -1.65685 0.0 0.0
   C17   O19      C20   H23   0.41421 1.24265 0.0 -1.65685 0.0 0.0
   H14   C13      C11   H12   30.334  0.0  -30.334 0.0  0.0  0.0
   H14   C13      C15   H16   30.334  0.0  -30.334 0.0  0.0  0.0

[ impropers ]
; ai   aj     ak    al       th0         cth
   C6    C8    C15    C7      0           334.72
   C15   C7    C8    C10      0           334.72
   C7    C8    C10   C11      0           334.72
   C8    C10   C11   C13      0           334.72
   C8    C7    C15   C13      0           334.72
   C10   C11   C13   C15      0           334.72
   C7    C15   C13   C11      0           334.72
   C13   C15   C11   H14      0           334.72
   H9    C7    C10   C8       0           334.72
   H12   C13   C10   C11      0           334.72
   H16   C13   C7    C15      0           334.72
   C17   C11    C8    C10     0           334.72
   C5   -C15   -C11 -C13      0           334.72
```

APPENDIX E

DIFFUSION CALCULATION

To calculate the diffusivity of water inside the mPPE channels, the mean square displacemens (MSD) of water molecules inside the mPPE channels in *Z* direction were monitored, and the diffusivity was calculated using Einstein's relation for the diffusivity. Specifically, the 1-D diffusivity along the *Z* axis, D_Z, was calculated using the following equation:

$$\lim_{t\to\infty}(\mathrm{MSD})_Z = 2D_Z t \tag{E.1}$$

where $(MSD)_Z$ is mean square displacement in *Z* direction and *t* is time.

The self diffusivity or three-dimensional diffusivity, D, of water molecules outside of the mPPE nano channels was calculated using the following equation:

$$\lim_{t\to\infty}\mathrm{MSD} = 6Dt \tag{E.2}$$

where MSD is three dimensional mean square displacement and *t* is time.

The following python script is used to monitor the position of all water molecules during simulation. It also calculates the MSD of each water molecule and the diffusivity of water.

E.1 Python script to calculate the diffusivity

```
#!/usr/bin/python

# Diffusion calculation script version 0.11
# Addtional features:
# 1. Automatically check for the length of the helix channel,
# based on the coordinates of atoms in the first and the last residues
# 2. Automatically check the time interval, reset the time interval for
# diffusion calculation if it's less than the frame time interval or
# greater than the total length of the trajectory.
# 3. Export data to three different ouput files:
#     - diffussion statistic file: diff_result.dat
#     - list of pore_water molecules in every time frame:
pore_water.dat
#     - list of pore_water molecules forming hydrogen bonds with
interior functional groups
#      and not move during the simulation
# 4. Calculate and export two diffusivities results of pore_water
#     - include all pore_water molecules
#     - exclude pore_water molecules stuck to the interior wall
#     (which did not move out of a sphere, centered by its coordinates
at the frame 0, r = 3.0A, in all frame)
#
# The diffusivity of pore water is calculated from Einstein equation
# based on the movement of oxygen atom of pore water (ok but better if
# using the center of mass).
# Pymacs must be installed for this scripts. Pymacs
# version 0.1, http://wwwuser.gwdg.de/~dseelig/pymacs.html. This
# module need to be installed on Palmetto.
# The trr trajectory need to be converted to an xtc trajectory
# use trjconv.



import sys
import csv
from geometry import geometry
from pymacs import *
from math import sqrt
from optparse import OptionParser

# Parse Input Options
parser = OptionParser()
parser.add_option("-f", dest="xtc",
                        help=" Input trajectory file (.xtc)",
                        metavar=" FILE",
                        default=None)
parser.add_option("-a", dest="atoms",
```

```
                        help=" List of water oxygen atoms",
                        metavar=" FILE",
                        default=None)
parser.add_option("-t", dest="time_interval",
                        help=" Time interval (ps)",
                        type="float",
                        metavar=" dTIME",
                        default=None)
parser.add_option("-b", dest="begin_time",
                                help=" Beginning at (ps)",
                                type="float",
                                metavar=" bTIME",
                                default=None)
parser.add_option("-e", dest="end_time",
                                help=" Ending at (ps)",
                                type="float",
                                metavar=" eTIME",
                                default=None)
parser.add_option("-r", dest="set_radius",
                                help=" distance from the opening",
                                type="float",
                                metavar=" RADIUS",
                                default=None)
parser.add_option("-d", dest="set_distance",
                                help=" distance from pore center",
                                type="float",
                                metavar=" CDISTANCE",
                                default=None)
parser.add_option("-v", dest="Catoms",
                                help=" List of 2 C (named C10) atoms 1
at residue 7, the other at residue 89",
                                metavar=" FILE",
                                default=None)
parser.add_option("-l", dest="C10layers",
                                help=" List of C10 of 8 helix sections,
6 C10 per section",
                                metavar=" FILE",
                                default=None)


(options, args) = parser.parse_args()
if None in [options.xtc, options.time_interval, options.begin_time,
options.end_time, options.atoms, options.set_radius,
options.set_distance, options.C10layers, options.Catoms]:
      parser.print_help()
      sys.exit()

# Class of water molecules inside helix cavity
# x0, x1, x2: coordinates of the oxygen atom in the current frame
# xc1, xc2, xc3: coordiantes of the oxygen atom in the first frame
# status: check if it's in pore: 0 if in; 1 if out
# f_status: check if it moves: 0 if move; 1 if not

class Pore_water:
```

```
      population=0

      def __init__(self, id, x0, x1, x2, xc0, xc1, xc2, tin, z_enter,
z_out, time, sdist, xydist, zdist, status, f_status):
            self.id           = id
            self.tin    = tin
            self.time   = time
            self.x0           = x0
            self.x1           = x1
            self.x2           = x2
            self.xc0    = xc0
            self.xc1    = xc1
            self.xc2    = xc2
            self.z_enter      = z_enter
            self.z_out  = z_out
            self.sdist  = sdist
            self.xydist = xydist
            self.zdist  = zdist
            self.status = status
            self.f_status     = f_status
            Pore_water.population += 1

#class of diffusion statistic data
class Aver_Dist:
      stat=0
      def __init__(self, sdist, xydist, zdist, time_com, count, diff1,
diff2, xydiff, zdiff):
            self.sdist  = sdist
            self.xydist = xydist
            self.zdist  = zdist
            self.time_com     = time_com
            self.count  = count
            self.diff1  = diff1
            self.diff2  = diff2
            self.xydiff = xydiff
            self.zdiff  = zdiff
            Aver_Dist.stat += 1

#Sphere of hydrogen bond
R_CHECK = 0.09 #(nm2, radius 3.0A)

#channel length: from the first residue + d_set to the last residue -
d_set
#the path length is now only 5.0 nm
D_SET = options.set_radius #(nm)
DISTANCE = options.set_distance #allowed distance from the channel
center
Z_DISTANCE = 0.4  #distance from the closest channel section
print "D_SET: distance from the opening:", D_SET
print "DISTANCE: distance from the center point:",DISTANCE
print "Z_DISTANCE: distance from the section center:", Z_DISTANCE

#number of water molecules go through the channel:
COUNT_WATER_THROUGH = 0
```

```
# Read oxygen atoms (water only) list. For pymacs 0.2, need to help the
list of oxygen water
# The diffusivity of pore_water is calculates based on the diffusivity
of oxygen water (for simplicity)
# For pymacs 0.4, this section could be replaced by model.atom.name

waters=open(options.atoms).readlines()

for i,water in enumerate(waters):
      waters[i]=water.split()
      for j,atm in enumerate(waters[i]):
            # XTC file object uses zero-based indices for atoms
            waters[i][j]=int(atm)-1

#read the C10 atoms at the beginning (residue 7) and at the end
(residue 89) of the helical polymer: number 1 and 2
catoms=open(options.Catoms).readlines()
for i,catom in enumerate(catoms):
      catoms[i]=catom.split()
      for j,atm in enumerate(catoms[i]):
            catoms[i][j]=int(atm)-1

#read the C15 atoms of the 8 sections of the helix channel: each
section has 6 C10 atoms
#section 1: C10 of residues 7 8 9 10 11 12
#section 2: C10 of residues 19 20 21 22 23 24
#section 3: C10 of residues 31 32 33 34 35 36
#section 4: C10 of residues 42 43 44 45 46 47
#section 5: C10 of residues 54 55 56 57 58 59
#section 6: C10 of residues 66 67 68 69 70 71
#section 7: C10 of residues 78 79 80 81 82 83
#section 8: C10 of residues 84 85 86 87 88 89
#ignore the first and the last helix layers

c10layers=open(options.C10layers).readlines()
for i,c10layer in enumerate(c10layers):
        c10layers[i]=c10layer.split()
        for j,atm in enumerate(c10layers[i]):
                c10layers[i][j]=int(atm)-1

#open the xtc file, find the time end
xtcfile=openXTC(options.xtc)
while True:

      # the coordinates are stored in frame['x']
        f = readXTCFrame(xtcfile)
        if not f:
                break
        frame=f
      TIME_END=int(frame['time'])

print "Total time of this xtc:", TIME_END
```

```
#Initiate set of water in helix pore, statistic
BTIME=int(options.begin_time)
ETIME=int(options.end_time)
if ETIME > TIME_END:
      ETIME = TIME_END


#initialize the diffusion data at dt time interval
dstat=[]
j=0
DTIME=int(options.time_interval)
for i in range (0,TIME_END,+DTIME):
      dstat.append([])
      dstat[j]=Aver_Dist(0, 0, 0, i, 0, 0, 0, 0, 0)
      j+=1

#Read the first frame in the .xtc file, find any water in the pore, add
to the list
pwater=[]

xtcfile=openXTC(options.xtc)
while True:

      # the coordinates are stored in frame['x']
      f = readXTCFrame(xtcfile)
      if not f:
            break
      frame=f
      #reading box dimensions
      B=[0,0,0]
        B[0]=frame['box'][0][0]
        B[1]=frame['box'][1][1]
        B[2]=frame['box'][2][2]
      print "Box dimensions:", B[0], B[1], B[2]
      max_d=B[0]
      MAX_ID=0
      for i in range(3):
            if max_d<B[i]:
                  max_d=B[i]
                      MAX_ID=i
      print "Longest dimension x:", MAX_ID, max_d

      #find the length of the helix cavity
      H_END=[]
      H_END1=[]
      i=0
      for catom in catoms:
            for atm in catom:
                  x=frame['x'][atm]
                  H_END.append([])
                  H_END1.append([])
                  H_END[i] = x[MAX_ID]
                  H_END1[i]=H_END[i]
```

```
                i+=1
        if H_END[0]>H_END[1]:
              temp=H_END[0]
              H_END[0]=H_END[1]
              H_END[1]=temp
        print "Helix from end to end:"
        print H_END[0], H_END[1]
        print "Helix in calculation:"
        print H_END[0] + D_SET, H_END[1]-D_SET

        # find the center for each helix section
        #print "C10 list:"
        x_c10=[]
        y_c10=[]
        z_c10=[]
        i=0
          for c10layer in c10layers:
                   for atm in c10layer:
                           x=frame['x'][atm]
                           x_c10.append([])
                     y_c10.append([])
                     z_c10.append([])
                           x_c10[i] = x[0]
                     y_c10[i] = x[1]
                     z_c10[i] = x[2]
           #                print '%5.2f   %5.2f   %5.2f' %(x_c10[i],
y_c10[i], z_c10[i])
                     i += 1
        center_x=[]
        center_y=[]
        center_z=[]
        #print "helix section center:"
        for i in range(0,8):
              temp_x = 0
              temp_y = 0
              temp_z = 0
              for j in range(0,6):
                    temp_j=i*6+j
                    temp_x = temp_x + x_c10[temp_j]/6
                    temp_y = temp_y + y_c10[temp_j]/6
                    temp_z = temp_z + z_c10[temp_j]/6
              center_x.append([])
              center_y.append([])
              center_z.append([])
              center_x[i] = temp_x
              center_y[i] = temp_y
              center_z[i] = temp_z
        #     print '%5.2f  %5.2f %5.2f' %(center_x[i], center_y[i],
center_z[i])
        break


xtcfile=openXTC(options.xtc)
while True:
```

```
        # the coordinates are stored in frame['x']
        f = readXTCFrame(xtcfile)
        if not f:
                break
        frame=f

        if (frame['time'] < BTIME):
                continue
      if BTIME>0:
            i=0
              for c10layer in c10layers:
                      for atm in c10layer:
                           x=frame['x'][atm]
                         x_c10[i] = x[0]
                               y_c10[i] = x[1]
                             z_c10[i] = x[2]
                             i+=1
              for i in range(0,8):
                    temp_x = 0
                  temp_y = 0
                      temp_z = 0
                    for j in range(0,6):
                        temp_j=i*6+j
                          temp_x = temp_x + x_c10[temp_j]/6
                              temp_y = temp_y + y_c10[temp_j]/6
                              temp_z = temp_z + z_c10[temp_j]/6
                  center_x[i] = temp_x
                    center_y[i] = temp_y
                  center_z[i] = temp_z

      #print "Waters in helix cavity (frame 0):"
      j=0
      for water in waters:
            for atm in water:
                  x=frame['x'][atm]
                  #test 1: check if the molecule within the z
coordinate
                  if (x[MAX_ID]>(H_END[0]+ D_SET)) and
(x[MAX_ID]<(H_END[1]- D_SET)):
                  #     print '%5i %5.2f  %5.2f %5.2f' %(atm, x[0],
x[1], x[2])
                  #find the two section closest to the above molecule:
                        position = 0
                        for i in range(0,8):
                              if abs(x[MAX_ID] > center_z[i]) <
Z_DISTANCE:
                                    position = i
                  #     print "Position: ", position
                        if (abs(x[0] - center_x[position])<DISTANCE)
and (abs(x[1] - center_y[position])<DISTANCE):
                              #add in the pore water list if sastify
                  #           print "OK"
                              pwater.append([])
```

```
                                pwater[j] = Pore_water(atm, x[0], x[1],
x[2], x[0], x[1], x[2], BTIME, x[MAX_ID], x[MAX_ID], 0, 0, 0, 0, 0, 1)
                                j += 1
      print "Total number of pore water:", Pore_water.population


      break


#Loop over all frames, check the water molecules in the helix cavity
(pwater list)
#If that water is still in the pore, calculate the distance it travels,
update the time
#If that water is nolonger in the pore, remove from the list

#print " Frame          No_water_remain         No_water_current
      No_water_in       No_water_out"

FILE_WATER=open('file_pore_water.dat',"w")
water_line=" Pore_water_id    Time_in           Time_stay" + "\n"
FILE_WATER.write(water_line)
FILE_WATER1=open('file_pore_water_through.dat',"w")
line="Pore_water_id     Time_in           Time stay   Z_enter
      Z_out" +"\n"
FILE_WATER1.write(line)
FILE_WATER2=open('file_pore_water_list.dat',"w")
line2=" Frame           No_water_remain         No_water_current
No_water_in             No_water_out"
FILE_WATER2.write(line2)

COUNT_WATER = int(Pore_water.population)
WATER_IN = 0
WATER_OUT = 0
WATER_REMAIN = 0
WATER_CURRENT = 0
COUNT_FRAME = 0
COUNT_UP = 0
COUNT_DOWN = 0
COUNT_UP_PASS = 0
COUNT_DOWN_PASS = 0


xtcfile=openXTC(options.xtc)
while True:

        # the coordinates are stored in frame['x']
        f = readXTCFrame(xtcfile)
        if not f:
                break
        frame=f
      if (frame['time'] < (BTIME+10)):
           continue

 #       print "Frame #:%10i:"%(frame['time'])
```

```
       #update the length of the helix cavity at the current frame
        i=0
        for catom in catoms:
                for atm in catom:
                        x=frame['x'][atm]
                        H_END1[i] = x[MAX_ID]
                        i+=1
        channel_length=abs(H_END1[1]-H_END1[0])

        #update the center of eight helix sections:
        i=0
        for c10layer in c10layers:
                for atm in c10layer:
                        x=frame['x'][atm]
                        x_c10[i] = x[0]
                        y_c10[i] = x[1]
                        z_c10[i] = x[2]
            #     print '%5.2f   %5.2f   %5.2f' %(x_c10[i], y_c10[i],
z_c10[i])
                        i=i+1
        for i in range(0,8):
                temp_x = 0
                temp_y = 0
                temp_z = 0
                for j in range(0,6):
                  temp_j=i*6+j
                        temp_x = temp_x + x_c10[temp_j]/6
                        temp_y = temp_y + y_c10[temp_j]/6
                        temp_z = temp_z + z_c10[temp_j]/6
                center_x[i] = temp_x
                center_y[i] = temp_y
                center_z[i] = temp_z
            #print '%5.2f  %5.2f %5.2f' %(center_x[i], center_y[i],
center_z[i])

      #1.calculate distance traveled by pore water
      #2.remove water molecules ouf of the helix cavity from the list
(status = 1)
      for i in range(Pore_water.population):
            x=frame['x'][pwater[i].id]
            if (x[MAX_ID]>(H_END[0] + D_SET)) and (x[MAX_ID]<(H_END[1]-
D_SET)) and (pwater[i].status==0):
            #     print '%5i  %5.2f  %5.2f   %5.2f' %(atm, x[0], x[1],
x[2])
                  position = 0
                        for j in range(0,8):
                             if abs(x[MAX_ID] -
center_z[j])<Z_DISTANCE:
                                position = j
                        if (abs(x[0]-center_x[position])<DISTANCE) and
(abs(x[1]-center_y[position])<DISTANCE):
                        pwater[i].sdist +=((x[0]-pwater[i].x0)**2 +
(x[1]-pwater[i].x1)**2 + (x[2]-pwater[i].x2)**2)
```

```
                        pwater[i].xydist += ((x[0]-pwater[i].x0)**2 +
(x[1]-pwater[i].x1)**2)
                        pwater[i].zdist += (x[2]-pwater[i].x2)**2
                        pwater[i].time=frame['time'] - pwater[i].tin
                        #reset the previous coordination to the new
coordination
                        pwater[i].x0 = x[0]
                        pwater[i].x1 = x[1]
                        pwater[i].x2 = x[2]

                  #calculate square distance from the original
coordinate at frame 0
                  #if larger than cut off radius r_check, f_status = 0,
and it will not be checked in the rest of the simulation
                  if pwater[i].f_status == 1:
                        temp = (x[0]-pwater[i].xc0)**2 + (x[1]-
pwater[i].xc1)**2 + (x[2]-pwater[i].xc2)**2
                        if temp > R_CHECK:
                              pwater[i].f_status = 0
            else:
                  pwater[i].status = 1
                  pwater[i].z_out = x[MAX_ID]

      # remove jumping water molecules
        # print the list of water in and water through
        temp=int(Pore_water.population)
      temp1=temp
      COUNT_OUT=0
      for j in range(temp1):
            for i in range(temp):
                  if pwater[i].status == 1:
                        if pwater[i].time ==0:
                              pwater.remove(pwater[i])
                              Pore_water.population -= 1
                              break
                        else:
                  #     check if this water move through the channel
                              if (abs(pwater[i].z_enter -
pwater[i].z_out) > 4):
                                    if pwater[i].z_out > 4:
                                          test=abs(pwater[i].z_out -
H_END[1] + D_SET)
                                    else:
                                          test=abs(pwater[i].z_out -
H_END[0] - D_SET)
                                    if test<1.0:
                                          COUNT_WATER_THROUGH =
COUNT_WATER_THROUGH + 1
                                          line=str(pwater[i].id) + "
" + str(pwater[i].tin) + "    " + str (pwater[i].time) + "  " +
str(pwater[i].z_enter)  + "  " + str(pwater[i].z_out) + "\n"
                                          FILE_WATER1.write(line)
                                          if pwater[i].z_enter > 4:
                                                COUNT_UP_PASS += 1
```

```
                                        else:
                                              COUNT_DOWN_PASS += 1

                                water_line=str(pwater[i].id) + "     " +
str(pwater[i].tin) + "    " + str (pwater[i].time) + "  " +
str(pwater[i].z_enter) + "  " + str(pwater[i].z_out) + "\n"
                                FILE_WATER.write(water_line)
                                pwater.remove(pwater[i])
                                COUNT_WATER += 1
                                COUNT_OUT += 1
                                Pore_water.population -= 1
                                break
            temp=int(Pore_water.population)
      water_remain=int(Pore_water.population)

      #add to the diffusion statistics
      for i in range(Aver_Dist.stat):
            for j in range(Pore_water.population):
                  if (pwater[j].time == dstat[i].time_com) and
(pwater[j].status==0):
                        dstat[i].sdist += pwater[j].sdist
                        dstat[i].xydist += pwater[j].xydist
                        dstat[i].zdist += pwater[j].zdist
                        dstat[i].count += 1

      #check if any new water going into the pore
      #print "New water molecules entering the helix cavity:"
      COUNT_NEW=0
      for water in waters:
                  for atm in water:
                        x=frame['x'][atm]
                  flag=0
                        if (x[MAX_ID]>(H_END[0] + D_SET)) and
(x[MAX_ID]<(H_END[1] - D_SET)):
                        #print '%5i  %5.2f  %5.2f   %5.2f' %(atm, x[0],
x[1], x[2])
                         pos = 0
                               for i in range(0,8):
                                if abs(x[MAX_ID] -
center_z[i])<Z_DISTANCE:
                                   pos = i
                        #print "position", position
                             if (abs(x[0]-center_x[pos])<DISTANCE) and
(abs(x[1]-center_y[pos])<DISTANCE):
                                for j in range(Pore_water.population):
                                      if (atm == pwater[j].id) and
(pwater[j].status == 0):
                                            flag=1
                                            break
                                if flag == 0:
                                      temp=Pore_water.population
                                      pwater.append([])
```

```
                                        pwater[temp] = Pore_water(atm,
x[0], x[1], x[2], x[0], x[1], x[2], frame['time'], x[MAX_ID],
x[MAX_ID], 0, 0, 0, 0, 0, 0)
                                        COUNT_NEW += 1
                                        if pwater[temp].z_enter > 5:
                                              COUNT_UP += 1
                                          else:
                                              COUNT_DOWN += 1


      water_current=int(Pore_water.population)
      WATER_IN += COUNT_NEW
      WATER_OUT += COUNT_OUT
      WATER_CURRENT += water_current
      WATER_REMAIN += water_remain
      COUNT_FRAME += 1

      #print "%12i %12i %12i %12i %12i %12i %12i %5.2f"
%(frame['time'], water_remain, water_current, COUNT_NEW, COUNT_UP,
COUNT_DOWN, COUNT_OUT, channel_length)
      line2 = str(frame['time']) + "  " + str(water_remain) + "  " +
str (water_current) + "  " + str(COUNT_NEW) + "  " + str(COUNT_OUT) + "
" + str(channel_length) + "\n"
      FILE_WATER2.write(line2)
      print "ID   x0    x1    x2    tin   time  sdist status
      f_status    z_enter     z_out"

      for i in range(Pore_water.population):
            if pwater[i].status==0:
                  print "%5i %5.2f  %5.2f %5.2f %5.2f %5.2f %5.2f %1.0f
%1.0f  %5.2f
%5.2f"%(pwater[i].id,pwater[i].x0,pwater[i].x1,pwater[i].x2,pwater[i].t
in,pwater[i].time,pwater[i].sdist,pwater[i].status, pwater[i].f_status,
pwater[i].z_enter, pwater[i].z_out)

# stop the loop after 10000ps
      if frame['time'] > ETIME:
            break

temp1 = float(WATER_IN)/float(COUNT_FRAME)
temp2 = float(WATER_OUT)/float(COUNT_FRAME)
temp3 = float(WATER_REMAIN)/float(COUNT_FRAME)
temp4 = float(WATER_CURRENT)/float(COUNT_FRAME)

print " Average/frame         WATER_IN    WATER_OUT   WATER_REMAIN
      WATER_CURRENT     No. Frame"
print " %8.3f   %8.3f   %8.3f   %8.3f %12i"     %(temp1, temp2, temp3,
temp4, COUNT_FRAME)
print "Total water counted in statistic"
print "     %8.3f %8.3f %8.3f       %8.3f %12i" %(COUNT_WATER,
WATER_OUT, WATER_REMAIN, WATER_CURRENT, COUNT_FRAME)
print "Total in         Up          Down        Total pass  Up    Down"
```

```
print " %12i   %12i   %12i    %12i   %12i    %12i" %(WATER_IN,
COUNT_UP, COUNT_DOWN, COUNT_WATER_THROUGH, COUNT_UP_PASS,
COUNT_DOWN_PASS)

line="Total number of water went through the channel: " +
str(COUNT_WATER_THROUGH) +"\n"
FILE_WATER1.write(line)
line="Total number of water ventured into the channel: " +
str(COUNT_WATER) + "\n"
FILE_WATER1.write(line)

FILE_WATER1.close()
FILE_WATER.close()
#FILE_WATER2.close()

line2 = " WATER_ID            TIME_IN                  TIME        SDIST
      STATUS            F_STATUS"
FILE_WATER2.write(line2)

print "Water molecules present in the helix cavity:"
for i in range(Pore_water.population):
        line2 = str(pwater[i].id) + "  " + str(pwater[i].tin) + "  " +
str(pwater[i].time) +  "  " + str(pwater[i].sdist) + "  " +
str(pwater[i].status) +  "  " + str(pwater[i].f_status) + "\n"
      FILE_WATER2.write(line2)

FILE_WATER2.close()


print "Diffusion statistic:"
print "(Time(ps)        Distance(nm**2)         Count(#of data points)
      Diffusion constant(nm**2/ps)  Diffusion(m**2/s) XY    Z"

for i in range(Aver_Dist.stat):
      if (dstat[i].count>0) and (dstat[i].time_com>0):
            dstat[i].diff1 =
dstat[i].sdist/((6*dstat[i].count)*(dstat[i].time_com))
            dstat[i].diff2 = 0.000001*dstat[i].diff1
            dstat[i].xydiff =
dstat[i].xydist/((4*dstat[i].count)*(dstat[i].time_com))
            dstat[i].zdiff =
dstat[i].zdist/((2*dstat[i].count)*(dstat[i].time_com))
            print "%12i %16.8f %12i %16.8f %18.12f %16.8f
%16.8f"%(dstat[i].time_com, dstat[i].sdist, dstat[i].count,
dstat[i].diff1, dstat[i].diff2, dstat[i].xydiff, dstat[i].zdiff)



#calculate the diffusion constant (at different time), exclude fixed
pore_water
#count the number of fixed pore_water molecules:

count_fixed = 0
```

```
print "Time(ps)         Distance(nm**2)         Count(#of data points)
      Diffusion constant(nm**2/ps)  Diffusion(m**2/s) XY    Z"

print "fixed water id:"
for i in range(Pore_water.population):
      if (pwater[i].status == 0) and (pwater[i].f_status == 1):
            count_fixed += 1
            print pwater[i].id

print "Total number of fixed water:", count_fixed
for i in range(Aver_Dist.stat):
      if (dstat[i].count>0) and (dstat[i].time_com>0):
            dstat[i].count = dstat[i].count - count_fixed
            dstat[i].diff1 =
dstat[i].sdist/((6*dstat[i].count)*(dstat[i].time_com))
            dstat[i].diff2 = 0.000001*dstat[i].diff1
            dstat[i].xydiff =
dstat[i].xydist/((4*dstat[i].count)*(dstat[i].time_com))
            dstat[i].zdiff =
dstat[i].zdist/((2*dstat[i].count)*(dstat[i].time_com))
            print "%12i %16.8f %12i %16.8f %18.12f %16.8f
%16.8f"%(dstat[i].time_com, dstat[i].sdist, dstat[i].count,
dstat[i].diff1, dstat[i].diff2, dstat[i].xydiff, dstat[i].zdiff)
```

Bash file to run the above python script

```
#!/bin/tcsh
#PBS -l nodes=1:ppn=1
#PBS -l walltime=01:10:00
#PBS -j oe
#PBS -M hhnguye@clemson.edu
#PBS -q main

module add intel/10.1
module add mpich/1.2.7
module add fftw
module add gromacs/4.0.5
setenv GMXLIB /home/hhnguye/top
setenv PYTHONPATH /projsmall/catalyst/share

#1. Convert the trajectories file .trr to .xtc
#2. Prepare the index file contain only aromatic carbon number C10 of
the
# polymer chain, name C10.ndx
#3. Prepare the index file contain oxygen atoms (OW) of all water
molecules
# name OW.ndx


cd $PBS_O_WORKDIR
```

```
set PATH='/projsmall/catalyst/hhnguye/channel/data/TPR'
set PATH1='/home/hhnguye/job/hhnguye/channel/run10'
set TEMP='channel_s513'
set FOLDER='/projsmall/catalyst/hhnguye/channel/data/OH'
set T1 = '4'
set T2 = '5'
set T3 = '6'
set ver = 'v25'

set COUNT=5

echo 'start on:'
date
cp diff_cal_${ver} ${T1}
sed -e
"s/file_pore_water.dat/diff_${TEMP}_${ver}_water_list_${COUNT}.dat/"
${T1} > ${T2}
sed -e
"s/file_pore_water_list.dat/diff_${TEMP}_${ver}_water_stat_${COUNT}.dat
/" ${T2} > ${T3}
sed -e
"s/file_pore_water_through.dat/diff_${TEMP}_${ver}_water_through_${COUN
T}.dat/" ${T3} > diff_cal_${TEMP}_${COUNT}
rm ${T1} ${T2} ${T3}
python  diff_cal_${TEMP}_${COUNT} -f ${TEMP}_final_cal.xtc -a
${TEMP}_OW.ndx -v ${TEMP}_C10.ndx -b 0 -e 100000 -t 20 -r 2.0 -d 1.2 -l
${TEMP}_C10_layers.ndx > diff_${TEMP}_${ver}_stat_${COUNT}.xvg
```

E.2 Modified version of the diffusion calculation

```
#!/usr/bin/python
# Diffusion calculation script version 0.11
# The diffusivity of pore water is calculated based on the flying time
# of the pore waters. Paper: van Hijkoop, V. J.; Dammers, A. J.; Malek,
# K.; Coppens, M. O., J. Chem. Phys. 2007, 127 (8), 085101(1-10)


import sys
import csv
from geometry import geometry
from pymacs import *
from math import sqrt
from optparse import OptionParser

# Parse Input Options
parser = OptionParser()
parser.add_option("-f", dest="xtc",
                        help=" Input trajectory file (.xtc)",
                        metavar=" FILE",
                        default=None)
```

```
parser.add_option("-a", dest="atoms",
                        help=" List of water oxygen atoms",
                        metavar=" FILE",
                        default=None)
parser.add_option("-t", dest="time_interval",
                        help=" Time interval (ps)",
                        type="float",
                        metavar=" dTIME",
                        default=None)
parser.add_option("-b", dest="begin_time",
                                help=" Beginning at (ps)",
                                type="float",
                                metavar=" bTIME",
                                default=None)
parser.add_option("-e", dest="end_time",
                                help=" Ending at (ps)",
                                type="float",
                                metavar=" eTIME",
                                default=None)
parser.add_option("-r", dest="set_radius",
                                help=" distance from the opening",
                                type="float",
                                metavar=" RADIUS",
                                default=None)
parser.add_option("-d", dest="set_distance",
                                help=" distance from pore center",
                                type="float",
                                metavar=" CDISTANCE",
                                default=None)
parser.add_option("-v", dest="Catoms",
                                help=" List of 2 C (named C10) atoms 1
at residue 7, the other at residue 89",
                                metavar=" FILE",
                                default=None)
parser.add_option("-l", dest="C10layers",
                                help=" List of C10 of 8 helix sections,
6 C10 per section",
                                metavar=" FILE",
                                default=None)


(options, args) = parser.parse_args()
if None in [options.xtc, options.time_interval, options.begin_time,
options.end_time, options.atoms, options.set_radius,
options.set_distance, options.C10layers, options.Catoms]:
      parser.print_help()
      sys.exit()

# Class of water molecules inside helix cavity
# x0, x1, x2: coordinates of the oxygen atom in the current frame
# xc1, xc2, xc3: coordiantes of the oxygen atom in the first frame
# status: check if it's in pore: 0 if in; 1 if out
# f_status: check if it moves: 0 if move; 1 if not
```

```
class Pore_water:
      population=0

      def __init__(self, id, x0, x1, x2, xc0, xc1, xc2, tin, z_enter,
z_out, time, sdist, xydist, zdist, status, f_status):
            self.id           = id
            self.tin    = tin
            self.time   = time
            self.x0           = x0
            self.x1           = x1
            self.x2           = x2
            self.xc0    = xc0
            self.xc1    = xc1
            self.xc2    = xc2
            self.z_enter      = z_enter
            self.z_out  = z_out
            self.sdist  = sdist
            self.xydist = xydist
            self.zdist  = zdist
            self.status = status
            self.f_status     = f_status
            Pore_water.population += 1

#class of diffusion statistic data
class Aver_Dist:
      stat=0
      def __init__(self, sdist, xydist, zdist, time_com, count, diff1,
diff2, xydiff, zdiff):
            self.sdist  = sdist
            self.xydist = xydist
            self.zdist  = zdist
            self.time_com     = time_com
            self.count  = count
            self.diff1  = diff1
            self.diff2  = diff2
            self.xydiff = xydiff
            self.zdiff  = zdiff
            Aver_Dist.stat += 1

#Sphere of hydrogen bond
R_CHECK = 0.09 #(nm2, radius 3.0A)

#channel length: from the first residue + d_set to the last residue -
d_set
#the path length is now only 5.0 nm
D_SET = options.set_radius #(nm)
DISTANCE = options.set_distance #allowed distance from the channel
center
Z_DISTANCE = 0.4  #distance from the closest channel section
print "OFFSET DISTANE:", D_SET
print "Center radius:", DISTANCE
print "Helix length in calculation: 4nm (2nm -- 6nm)"

#number of water molecules go through the channel:
```

```
COUNT_WATER_THROUGH = 0


# Read oxygen atoms (water only) list. For pymacs 0.2, need to help the
list of oxygen water
# The diffusivity of pore_water is calculates based on the diffusivity
of oxygen water (for simplicity)
# For pymacs 0.4, this section could be replaced by model.atom.name

waters=open(options.atoms).readlines()

for i,water in enumerate(waters):
      waters[i]=water.split()
      for j,atm in enumerate(waters[i]):
            # XTC file object uses zero-based indices for atoms
            waters[i][j]=int(atm)-1

#read the C10 atoms at the beginning (residue 7) and at the end
(residue 89) of the helical polymer: number 1 and 2
catoms=open(options.Catoms).readlines()
for i,catom in enumerate(catoms):
      catoms[i]=catom.split()
      for j,atm in enumerate(catoms[i]):
            catoms[i][j]=int(atm)-1

#read the C15 atoms of the 8 sections of the helix channel: each
section has 6 C10 atoms
#section 1: C10 of residues 7 8 9 10 11 12
#section 2: C10 of residues 19 20 21 22 23 24
#section 3: C10 of residues 31 32 33 34 35 36
#section 4: C10 of residues 42 43 44 45 46 47
#section 5: C10 of residues 54 55 56 57 58 59
#section 6: C10 of residues  66 67 68 69 70 71
#section 7: C10 of residues 78 79 80 81 82 83
#section 8: C10 of residues 84 85 86 87 88 89
#ignore the first and the last helix layers

c10layers=open(options.C10layers).readlines()
for i,c10layer in enumerate(c10layers):
        c10layers[i]=c10layer.split()
        for j,atm in enumerate(c10layers[i]):
                c10layers[i][j]=int(atm)-1

#open the xtc file, find the time end
xtcfile=openXTC(options.xtc)
while True:

      # the coordinates are stored in frame['x']
        f = readXTCFrame(xtcfile)
        if not f:
                break
        frame=f
      TIME_END=int(frame['time'])
```

```
print "Total time of this xtc:", TIME_END

#Initiate set of water in helix pore, statistic
BTIME=int(options.begin_time)
ETIME=int(options.end_time)
if ETIME > TIME_END:
      ETIME = TIME_END


#initialize the diffusion data at dt time interval
dstat=[]
j=0
DTIME=int(options.time_interval)
for i in range (0,TIME_END,+DTIME):
      dstat.append([])
      dstat[j]=Aver_Dist(0, 0, 0, i, 0, 0, 0, 0, 0)
      j+=1

#Read the first frame in the .xtc file, find any water in the pore, add
to the list
pwater=[]

xtcfile=openXTC(options.xtc)
while True:

      # the coordinates are stored in frame['x']
      f = readXTCFrame(xtcfile)
      if not f:
            break
      frame=f
      #reading box dimensions
      B=[0,0,0]
        B[0]=frame['box'][0][0]
        B[1]=frame['box'][1][1]
        B[2]=frame['box'][2][2]
      print "Box dimensions:", B[0], B[1], B[2]
      max_d=B[0]
      MAX_ID=0
      for i in range(3):
            if max_d<B[i]:
                  max_d=B[i]
                      MAX_ID=i
      print "Longest dimension x:", MAX_ID, max_d

      #find the length of the helix cavity
      H_END=[]
      H_END1=[]
      i=0
      for catom in catoms:
            for atm in catom:
                  x=frame['x'][atm]
                  H_END.append([])
                  H_END1.append([])
                  H_END[i] = x[MAX_ID]
```

```
                H_END1[i]=H_END[i]
                i+=1
        if H_END[0]>H_END[1]:
              temp=H_END[0]
              H_END[0]=H_END[1]
              H_END[1]=temp
        print H_END[0], H_END[1]

        # find the center for each helix section
        #print "C10 list:"
        x_c10=[]
        y_c10=[]
        z_c10=[]
        i=0
          for c10layer in c10layers:
                   for atm in c10layer:
                           x=frame['x'][atm]
                           x_c10.append([])
                     y_c10.append([])
                     z_c10.append([])
                           x_c10[i] = x[0]
                     y_c10[i] = x[1]
                     z_c10[i] = x[2]
           #               print '%5.2f   %5.2f   %5.2f' %(x_c10[i],
y_c10[i], z_c10[i])
                     i += 1
        center_x=[]
        center_y=[]
        center_z=[]
        #print "helix section center:"
        for i in range(0,8):
              temp_x = 0
              temp_y = 0
              temp_z = 0
              for j in range(0,6):
                    temp_j=i*6+j
                    temp_x = temp_x + x_c10[temp_j]/6
                    temp_y = temp_y + y_c10[temp_j]/6
                    temp_z = temp_z + z_c10[temp_j]/6
              center_x.append([])
              center_y.append([])
              center_z.append([])
              center_x[i] = temp_x
              center_y[i] = temp_y
              center_z[i] = temp_z
        #     print '%5.2f  %5.2f %5.2f' %(center_x[i], center_y[i],
center_z[i])
        break


xtcfile=openXTC(options.xtc)
while True:
          # the coordinates are stored in frame['x']
          f = readXTCFrame(xtcfile)
```

```
        if not f:
                break
        frame=f

        if (frame['time'] < BTIME):
                continue
      if BTIME>0:
            i=0
              for c10layer in c10layers:
                      for atm in c10layer:
                           x=frame['x'][atm]
                         x_c10[i] = x[0]
                               y_c10[i] = x[1]
                             z_c10[i] = x[2]
                             i+=1
              for i in range(0,8):
                    temp_x = 0
                  temp_y = 0
                      temp_z = 0
                    for j in range(0,6):
                        temp_j=i*6+j
                          temp_x = temp_x + x_c10[temp_j]/6
                              temp_y = temp_y + y_c10[temp_j]/6
                              temp_z = temp_z + z_c10[temp_j]/6
                  center_x[i] = temp_x
                    center_y[i] = temp_y
                  center_z[i] = temp_z

      #print "Waters in helix cavity (frame 0):"
      j=0
      for water in waters:
            for atm in water:
                  x=frame['x'][atm]
                  #test 1: check if the molecule within the z
coordinate
                  if (x[MAX_ID]> D_SET) and (x[MAX_ID]< D_SET + 4):
                  #     print '%5i %5.2f  %5.2f %5.2f' %(atm, x[0],
x[1], x[2])
                  #find the two section closest to the above molecule:
                        position = 0
                        for i in range(0,8):
                              if abs(x[MAX_ID] >
center_z[i])<Z_DISTANCE:
                                    position = i
                  #     print "Position: ", position
                        if (abs(x[0] - center_x[position])<DISTANCE)
and (abs(x[1] - center_y[position])<DISTANCE):
                              #add in the pore water list if sastify
                  #           print "OK"
                              pwater.append([])
                              pwater[j] = Pore_water(atm, x[0], x[1],
x[2], x[0], x[1], x[2], BTIME, x[MAX_ID], x[MAX_ID], 0, 0, 0, 0, 0, 1)
                              j += 1
      print "Total number of pore water:", Pore_water.population
```

```
      break


#Loop over all frames, check the water molecules in the helix cavity
(pwater list)
#If that water is still in the pore, calculate the distance it travels,
update the time
#If that water is nolonger in the pore, remove from the list

#print " Frame          No_water_remain         No_water_current
      No_water_in       No_water_out"

FILE_WATER=open('file_pore_water.dat',"w")
water_line=" Pore_water_id    Time_in           Time_stay   z_enter
      z_out" + "\n"
FILE_WATER.write(water_line)


xtcfile=openXTC(options.xtc)
while True:

        # the coordinates are stored in frame['x']
        f = readXTCFrame(xtcfile)
        if not f:
                break
        frame=f
      if (frame['time'] < (BTIME+10)):
            continue

 #       print "Frame #:%10i:"%(frame['time'])
       #update the length of the helix cavity at the current frame
        i=0
        for catom in catoms:
                for atm in catom:
                        x=frame['x'][atm]
                        H_END1[i] = x[MAX_ID]
                        i+=1
        channel_length=abs(H_END1[1]-H_END1[0])

        #update the center of eight helix sections:
        i=0
        for c10layer in c10layers:
                for atm in c10layer:
                        x=frame['x'][atm]
                        x_c10[i] = x[0]
                        y_c10[i] = x[1]
                        z_c10[i] = x[2]
            #     print '%5.2f   %5.2f   %5.2f' %(x_c10[i], y_c10[i],
z_c10[i])
                        i=i+1
        for i in range(0,8):
                temp_x = 0
```

```
                temp_y = 0
                temp_z = 0
                for j in range(0,6):
                  temp_j=i*6+j
                        temp_x = temp_x + x_c10[temp_j]/6
                        temp_y = temp_y + y_c10[temp_j]/6
                        temp_z = temp_z + z_c10[temp_j]/6
                center_x[i] = temp_x
                center_y[i] = temp_y
                center_z[i] = temp_z
            #print '%5.2f  %5.2f %5.2f' %(center_x[i], center_y[i],
center_z[i])

      #1.calculate distance traveled by pore water
      #2.remove water molecules out of the helix cavity from the list
(status = 1)
      for i in range(Pore_water.population):
            x=frame['x'][pwater[i].id]
            if (x[MAX_ID]>D_SET) and (x[MAX_ID]<D_SET + 4) and
(pwater[i].status==0):
            #     print '%5i  %5.2f  %5.2f   %5.2f' %(atm, x[0], x[1],
x[2])
                  position = 0
                        for j in range(0,8):
                             if abs(x[MAX_ID] -
center_z[j])<Z_DISTANCE:
                                position = j
                        if (abs(x[0]-center_x[position])<DISTANCE) and
(abs(x[1]-center_y[position])<DISTANCE):
                        pwater[i].sdist +=((x[0]-pwater[i].x0)**2 +
(x[1]-pwater[i].x1)**2 + (x[2]-pwater[i].x2)**2)
                        pwater[i].xydist += ((x[0]-pwater[i].x0)**2 +
(x[1]-pwater[i].x1)**2)
                        pwater[i].zdist += (x[2]-pwater[i].x2)**2
                        pwater[i].time=frame['time'] - pwater[i].tin
                        #reset the previous coordination to the new
coordination
                        pwater[i].x0 = x[0]
                        pwater[i].x1 = x[1]
                        pwater[i].x2 = x[2]

                  #calculate square distance from the original
coordinate at frame 0
                  #if larger than cut off radius r_check, f_status = 0,
and it will not be checked in the rest of the simulation
                  if pwater[i].f_status == 1:
                        temp = (x[0]-pwater[i].xc0)**2 + (x[1]-
pwater[i].xc1)**2 + (x[2]-pwater[i].xc2)**2
                        if temp > R_CHECK:
                              pwater[i].f_status = 0
            else:
                  pwater[i].status = 1
                  pwater[i].z_out = x[MAX_ID]
```

```
      # remove jumping water molecules
        # print the list of water in and water through
        temp=int(Pore_water.population)
      temp1=temp
      COUNT_OUT=0
      for j in range(temp1):
            for i in range(temp):
                  if pwater[i].status == 1:
                        if pwater[i].time ==0:
                              pwater.remove(pwater[i])
                              Pore_water.population -= 1
                              break
                        else:
                              water_line=str(pwater[i].id) + "     " +
str(pwater[i].tin) + "    " + str (pwater[i].time) + "  " +
str(pwater[i].z_enter) + "  " + str(pwater[i].z_out) + "\n"
                              FILE_WATER.write(water_line)
                              pwater.remove(pwater[i])
                              Pore_water.population -= 1
                              break
            temp=int(Pore_water.population)
      water_remain=int(Pore_water.population)

# stop the loop after 10000ps
      if (frame['time'] > ETIME) or (Pore_water.population == 0):
            break
```

Bash script to run the above python script

```
#!/bin/tcsh
#PBS -l nodes=1:ppn=1
#PBS -l walltime=01:00:00
#PBS -j oe
#PBS -M hhnguye@clemson.edu
#PBS -q main

module add intel/10.1
module add mpich/1.2.7
module add fftw
module add gromacs/4.0.5
setenv GMXLIB /home/hhnguye/top
setenv PYTHONPATH /projsmall/catalyst/share

cd $PBS_O_WORKDIR
set PATH='/projsmall/catalyst/hhnguye/channel/data/TPR'
set PATH1='/home/hhnguye/job/hhnguye/channel/run10'
set TEMP='channel_s513'
set FOLDER='/projsmall/catalyst/hhnguye/channel/data/OH'
set T1 = '4'
set T2 = '5'
```

```
set T3 = '6'

echo 'start at:'
date

set INTERVAL = 5000
set maxrun = 10
foreach COUNT (`seq 0 $maxrun`)
        set BEGIN_TIME=`echo "print '%8i'%($COUNT.0*$INTERVAL)"|python`
#        set END_TIME=`echo "print '%8i'%($BEGIN_TIME +
$INTERVAL.0)"|python`
        echo $COUNT
        echo $BEGIN_TIME
#        echo $END_TIME
      cp diff_cal_fly_time ${T1}
      sed -e
"s/file_pore_water.dat/diff_${TEMP}_water_time_exit_${COUNT}.dat/"
${T1} > ${T2}
      rm ${T1}
      python  ${T2} -f ${TEMP}_final_cal.xtc -a ${TEMP}_OW.ndx -v
${TEMP}_C10.ndx -b ${BEGIN_TIME} -e 100000 -t 20 -r 2.0 -d 1.2 -l
${TEMP}_C10_layers.ndx
      rm ${T2}
end

date
-----------------------------------------
```

APPENDIX F

CHEMICAL NUCLEAR MAGNETIC RESONANCE DATA

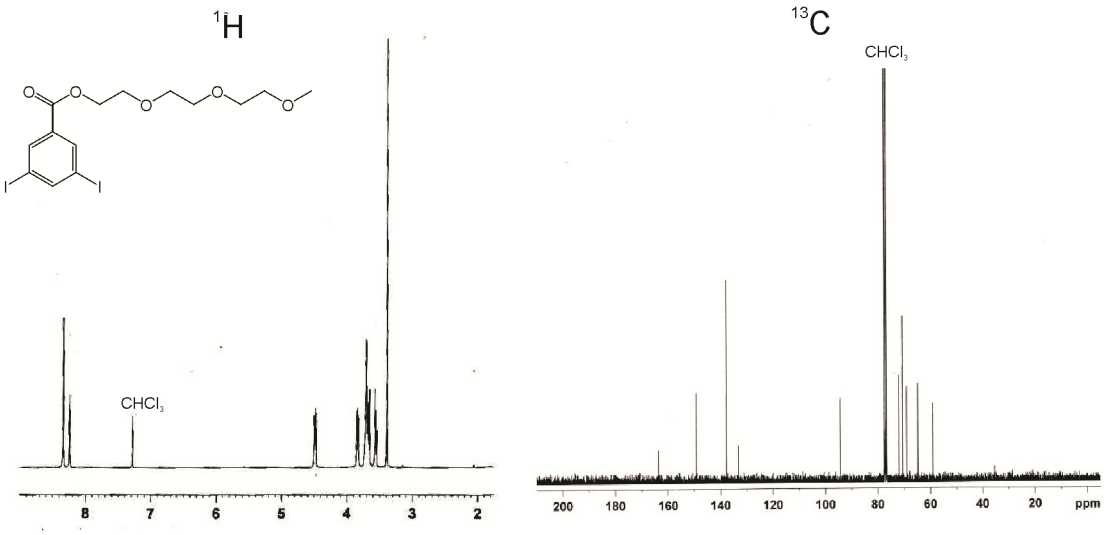


Figure F.1 NMR spectra for chemical (1).

^{1}H NMR (300MHz, $CDCl_3$, δ): 8.34 (d, J = 1.6 Hz, Ar H, 2H), 8.24 (t, J = 1.6 Hz, Ar H, 1H), 4.47 (t, J = 4.7 Hz, -COO-C*H*$_2$-CH_2-, 2H), 3.84 (t, J = 4.7 Hz, -COO-CH_2-C*H*$_2$-O-, 2H), 3.65-3.72 (m, -O-C*H*$_2$– C*H*$_2$-O-C*H*$_2$-CH_2-O-CH_3, 6H), 3.54-3.57 (m, –C*H*$_2$-O-CH_3, 2H), 3.39 (s, 3H, -CH_2-OC*H*$_3$); ^{13}C NMR (300MHz, $CDCl_3$, δ): 163.54, 149.16, 137.68, 133.24, 94.27, 71.87, 70.59, 70.53, 68.92, 64.67, 58.95.

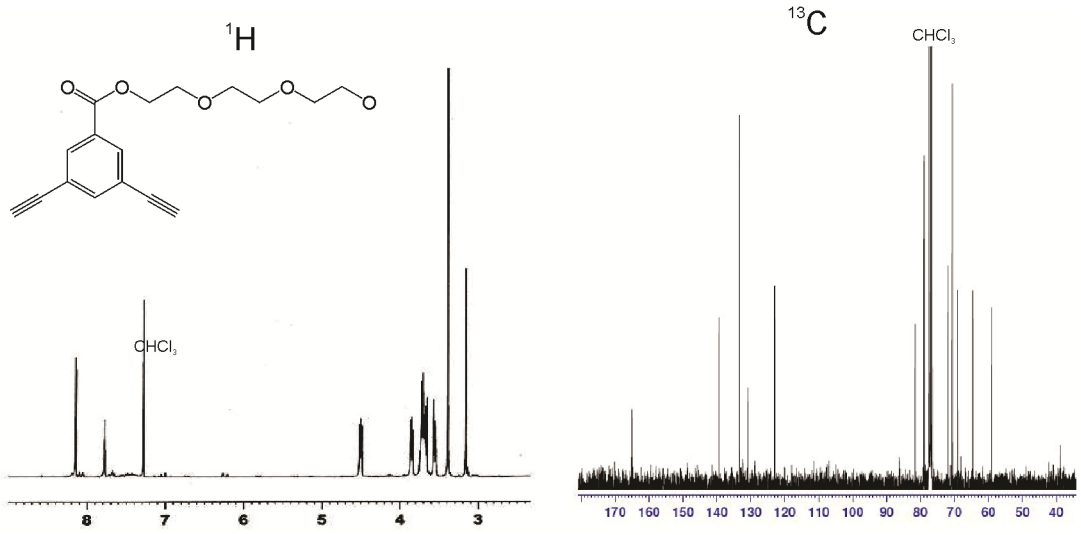


Figure F.2 NMR spectra for chemical (2).

^{1}H NMR (300MHz, $CDCl_3$, δ): 8.14 (s, Ar H, 2H), 7.77 (s, Ar H, 1H), 4.49 (t, J = 4.7 Hz, -COOC*H*$_2$-CH_2-, 2H), 3.85 (t, J = 4.7 Hz, -$COOCH_2$-C*H*$_2$-O-, 2H), 3.65-3.74 (m, -O-C*H*$_2$– C*H*$_2$-O-C*H*$_2$-CH_2-O-CH_3,6H), 3.54-3.57 (m, –C*H*$_2$-O-CH_3, 2H), 3.38 (s, -CH_2-OC*H*$_3$, 2H), 3.17 (s, -C≡C*H*, 2H). ^{13}C NMR (300MHz, $CDCl_3$,δ): 164.93, 139.38, 133.32, 130.83, 122.99, 81.79, 78.96, 71.93, 70.70, 70.63, 69.06, 64.58, 59.04.

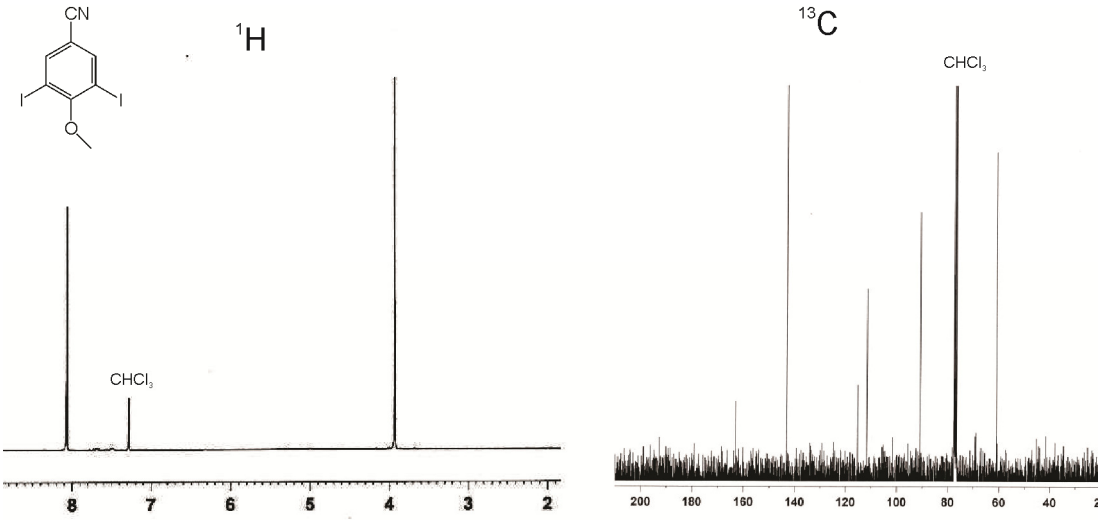


Figure F.3 NMR spectra for chemical (3).

^{1}H NMR (300MHz, $CDCl_3$, δ): 8.10 (s, Ar H, 2H), 3.95 (s, Ar-OC*H*$_3$, 3H); ^{13}C NMR (300MHz, $CDCl_3$, δ): 163.09, 143.11, 115.31, 111.65, 90.77, 60.95.

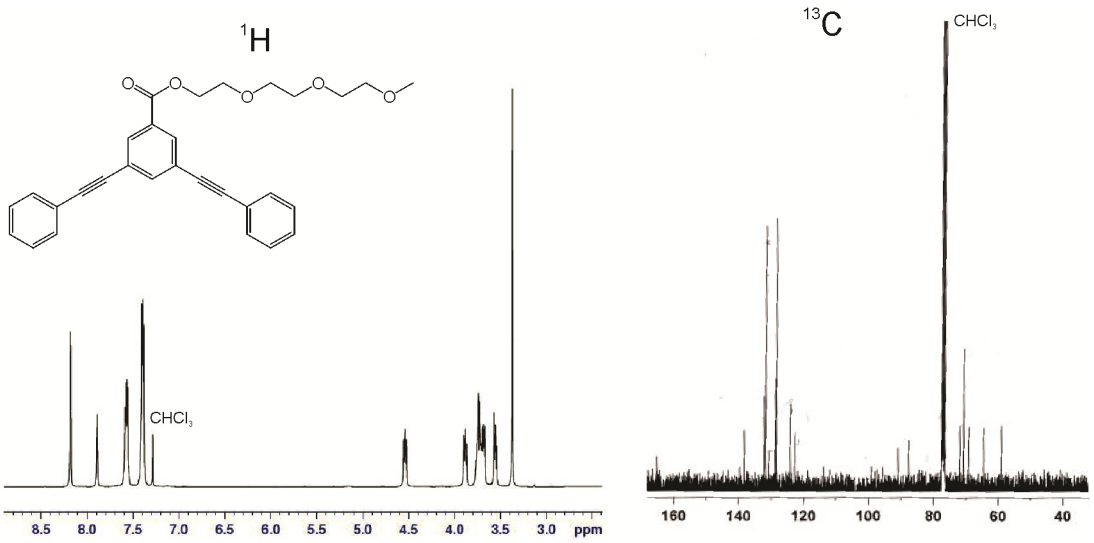


Figure F.4 NMR spectra for chemical (4) (mPPEa).

^{1}H NMR (300MHz, $CDCl_3$, δ): 8.18 (d, J = 1.4 Hz, Ar H, 2H), 7.88 (m, J = 1.4 Hz, Ar H, 2H), 7.55-7.59 (m, Ar H, 4H), 7.38-7.40 (m, Ar H, 6H), 4.54 (t, J = 4.7 Hz, -COOC*H*$_2$-, 2H), 3.89 (t, J = 4.7 Hz, -COOCH$_2$-C*H*$_2$-O-, 2H), 3.67-3.75 (m, -O-C*H*$_2$–C*H*$_2$-O-C*H*$_2$-CH$_2$-O-CH$_3$, 6H), 3.54-3.57 (m, –C*H*$_2$-O-CH$_3$, 2H), 3.38 (s, -OC*H*$_3$, 3H).^{13}C NMR (300MHz, $CDCl_3$, δ): 138.5, 132.2, 131.8, 130.9, 128.8, 128.5, 124.1, 90.9, 87.6, 71.93, 70.7, 70.6, 69.1, 64.6, 59.0.

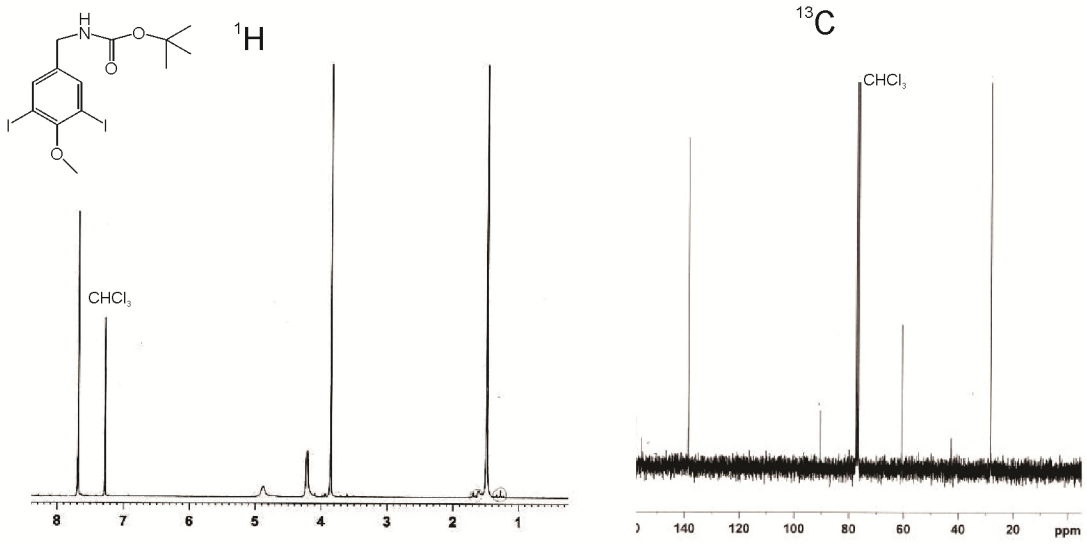


Figure F.5 NMR spectra for chemical (6).

^{1}H NMR (300MHz, $CDCl_3$, δ):7.68 (s, Ar H, 2H), 4.88 (br s, NH), 4.21 (d, J = 5.4 Hz, -C*H*$_2$-NH-, 2H), 3.85 (s, Ar-OC*H*$_3$, -3H,), 1.45 (s, C-(C*H*$_3$)$_3$, 9H); ^{13}C NMR (300MHz, $CDCl_3$, δ): 157.5, 156.1, 138.6, 91.4, 60.7, 43, 28.4.

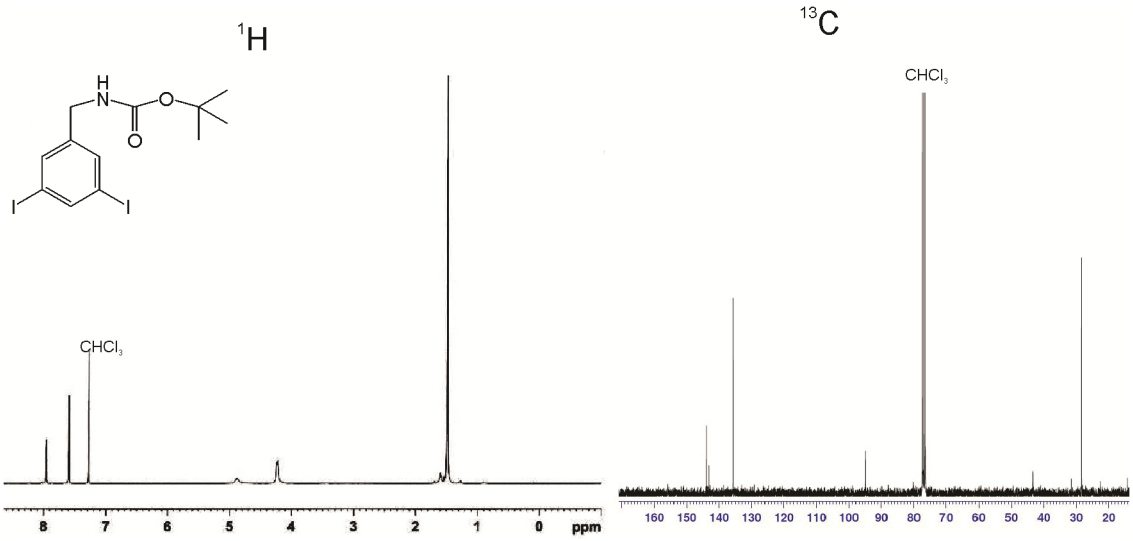


Figure F.6 NMR spectra for chemical (7).

^{1}H NMR 300MHz, $CDCl_3$, δ): 7.96 (t, J = 1.5 Hz, Ar H, 1H), 7.59-7.60 (m, Ar H, 2H), 4.90 (br s, -NH-), 4.22 (d, J = 5.5 Hz, -C*H*$_2$-NH-, 2H), 1.48 (s, -C(C*H*$_3$)$_3$, 9H); ^{13}C NMR (300MHz, $CDCl_3$, δ):146.75, 143.98, 143.20, 135.60, 94.9, 43.2, 28.36.

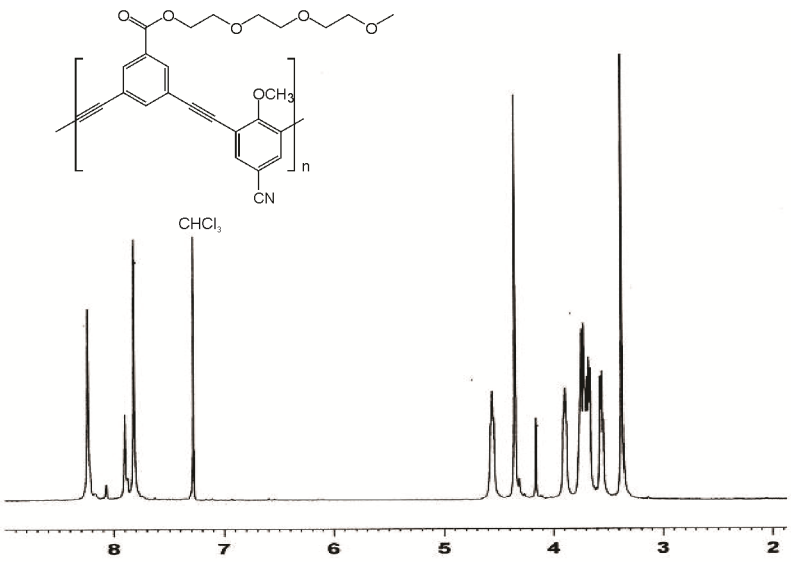


Figure F.7 1H NMR spectrum for mPPE1.

^{1}H NMR (300MHz, $CDCl_3$, δ): 8.24 (br s, Ar H), 7.90 (br s, Ar H, 1H), 7.81 (br s, Ar H), 4.54-4.57 (br m, -COOC*H*$_2$-CH_2-), 4.35 (br s, Ar-OC*H*$_3$), 3.88-3.91 (br m , -$COOCH_2$-C*H*$_2$-O-, 2H), 3.66-3.74 (br m, -O-C*H*$_2$–C*H*$_2$-O-C*H*$_2$-CH_2-O-CH_3), 3.55-3.57 (br m,-C*H*$_2$-O-CH_3), 3.37 (br s, -CH_2-OC*H*$_3$).

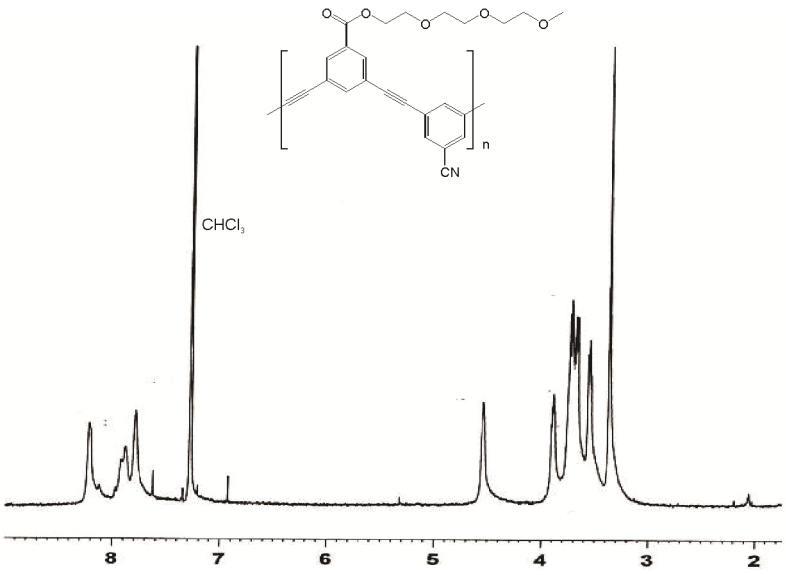


Figure F.8 1H NMR spectrum for mPPE2.

^{1}H NMR (300MHz, $CDCl_3$, δ)**:** 8.22 (br s, Ar H), 7.92 (br s, Ar H) 7.88 (br, Ar H), 7.79 (br, Ar H), 4.55 (br s, -COOCH_2-), 3.89 (br s, -CH_2-O-), 3.67-3.75 (br m, -O-CH_2–CH_2-O-CH_2-CH_2-O-CH_3), 3.54-3.57 br (m, -CH_2-O-CH_3), 3.37 (br s, -CH_2OCH_3).

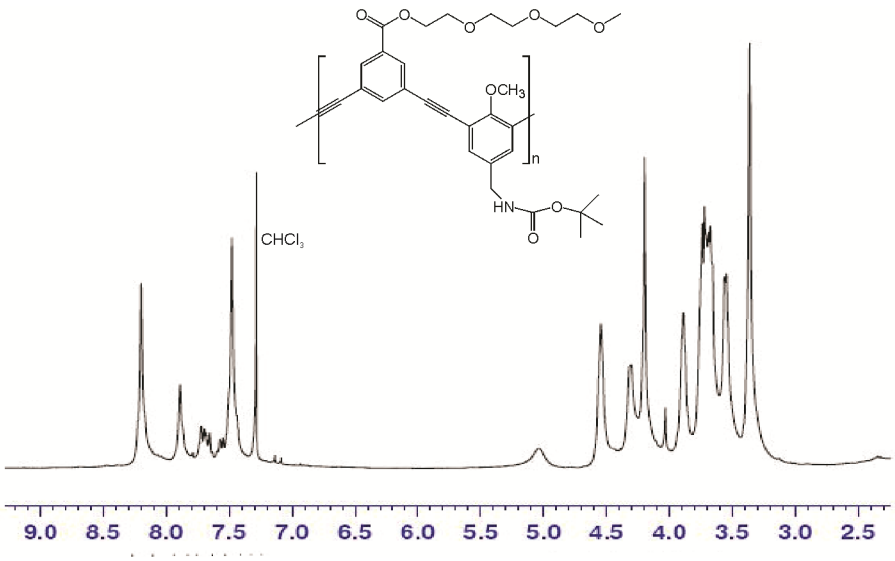


Figure F.9 1H NMR spectrum for p_mPPE3.

^{1}H NMR (300MHz, $CDCl_3$, δ): 8.19 (br s, Ar H), 7.88 (br s, Ar H**),** 7.47 (br s, Ar H, 2H), 5.01 (br s, -NH), 4.53-4.56 (br m, -COOCH_2), 4.23 (br s, -CH_2-NH-), 4.19 (br s, -OCH_3), 3.87-3.90 (br m, $COOCH_2$-CH_2-O-, 2H), 3.65-3.74 (br m, -O-CH_2–CH_2-O-CH_2-CH_2-O-CH_3), 3.55-3.58 (br m, -CH_2-O-CH_3), 3.37 (br s, CH_2-O-CH_3), 1.49 (br s, -C(CH_3)$_3$).

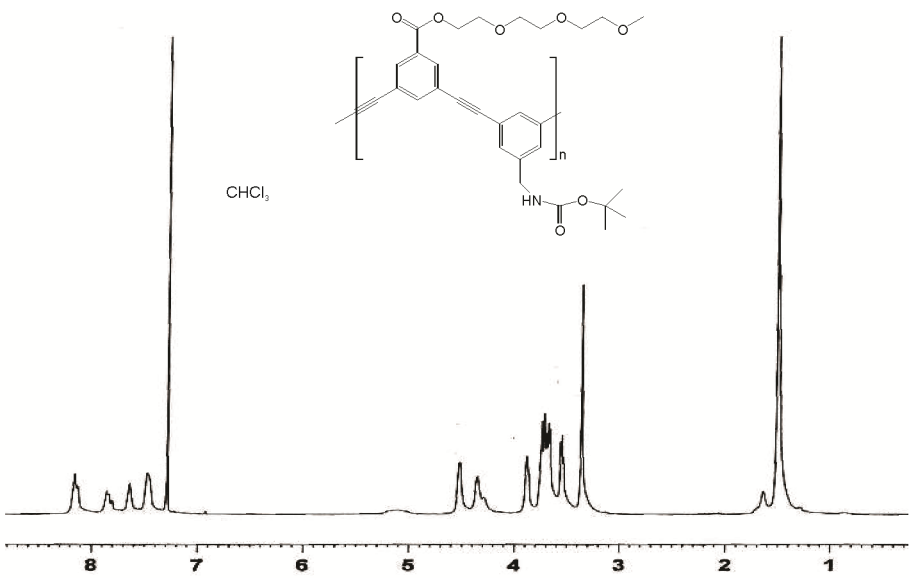


Figure F.10 1H NMR spectrum for p_mPPE4.

^{1}H NMR (300MHz, $CDCl_3$, δ): 8.13-8.15 (br m, Ar H), 7.80-7.85(br m, Ar H), 7.63 (br s, Ar H), 7.47 (br s, Ar H), 5.10 (br s, -NH), 4.52 (br s, -COOCH_2-), 4.35 (br s, -CH_2-NH-), 3.88 (br s, -COO-CH_2-CH_2-O-), 3.66-3.74 (m, -O-CH_2–CH_2-O-CH_2-CH_2-O-CH_3), 3.53-3.56 (m, -CH_2-O-CH_3), 3.36 (s, -CH_2-OCH_3), 1.51 (s, -C(CH_3)$_3$).

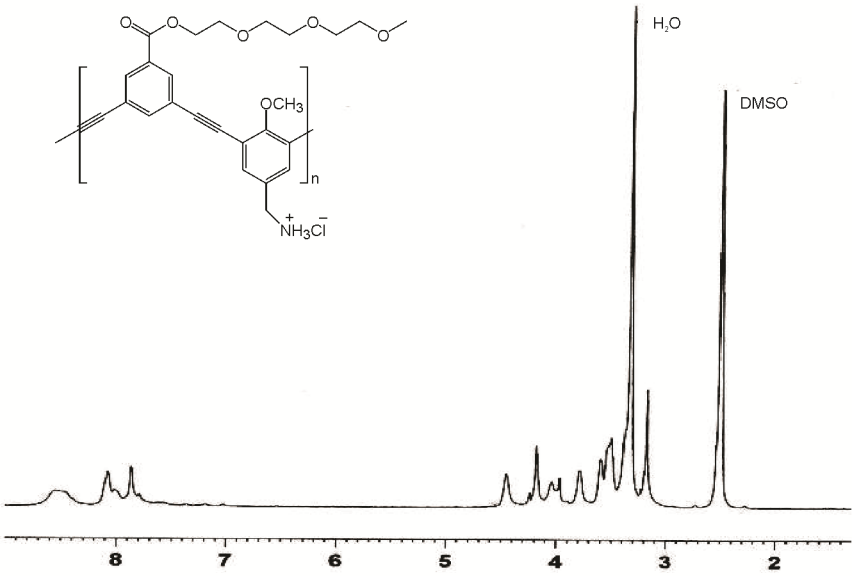


Figure F.11 1H spectrum for mPPE3.

^{1}H NMR (300MHz, DMSO-d6, δ): 8.54 (br s, -NH_3+), 8.02-8.08 (br m, Ar H), 7.79-7.87 (br m, Ar H), 4.45 (br s, -COOCH_2-), 4.18 (br s, -OCH_3), 4.04 (br s, -CH_2-NH-), 3.78-3.97 (br m, -CH_2-O-CH_2-), 3.50-3.60 (br m, -CH_2-O-CH_2-), 3.17 (s, -OCH_3).

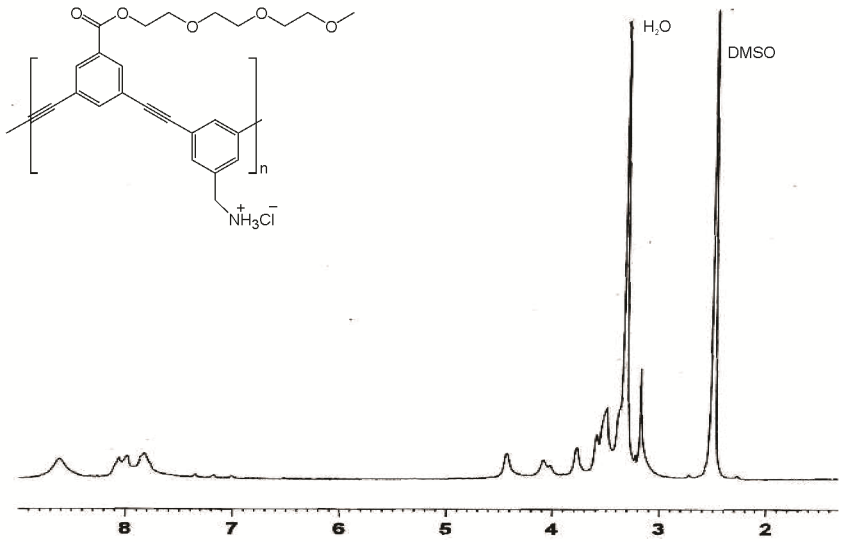


Figure F.12 1H NMR spectrum for mPPE4.

^{1}H NMR (300MHz, DMSO-d6, δ): 8.62 (br s, -NH_3+), 7.83-8.26 (br m, Ar H), 4.43 (br s, -COOCH_2-), 4.08-4.12 (br m, -CH_2-NH-), 3.77 (br s, -CH_2-O-CH_2-), 3.50-3.62(br m, -CH_2-O-CH_2-), 3.18 (s, -OCH_3).

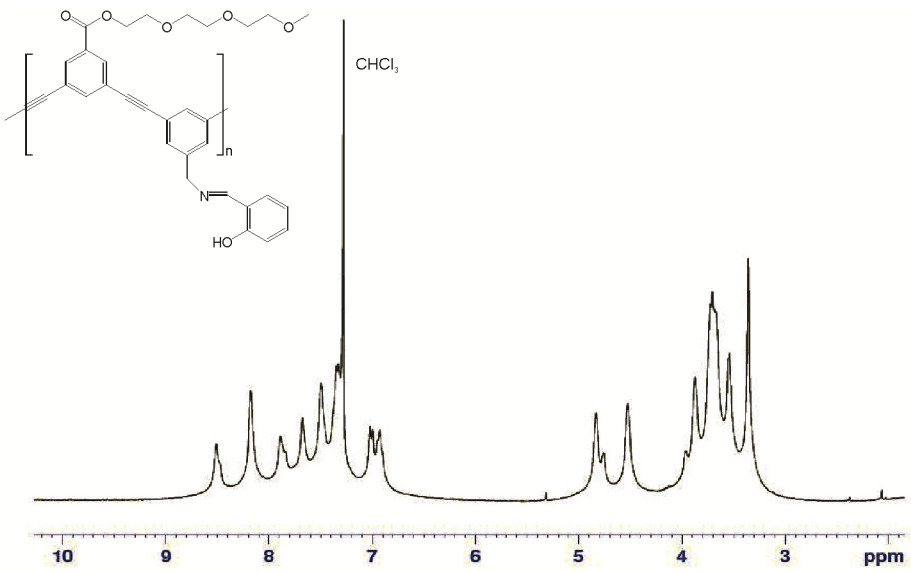


Figure F.13 1H NMR spectrum for imine functionalized mPPE.

^{1}H NMR (300MHz, $CDCl_3$, δ): 8.51(br s, N=C*H*), 8.17 (br s, Ar H), 7.84 (br s, Ar H), 7.67 (br s, Ar H), 7.49 (br s, Ar H), 7.32 (br m, Ar H), 7.02 (br m, Ar H), 6.94 (br m, Ar H), 4.83 (br s, -CH_2-N=C), 4.52 (br s, -COOCH_2-), 3.87(br s, -COOCH_2-CH_2-O-,), 3.66-3.72 (br m, -CH_2-CH_2-O-), 3.54 (br s, -CH_2-O-CH_3), 3.36 (br s, -OCH_3).

APPENDIX G

CHEMICAL FOURIER TRANSFORM INFARED SPECTRA

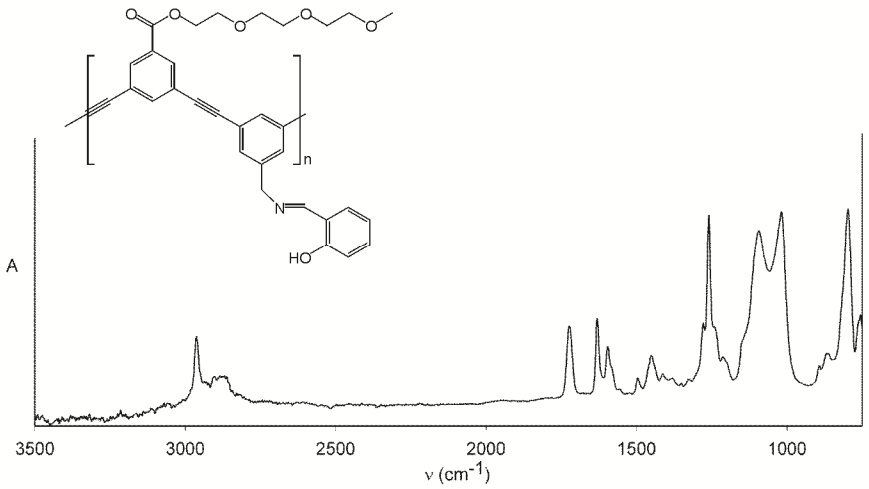


Figure G.1. FT-IR spectrum for imine functionalized mPPE.

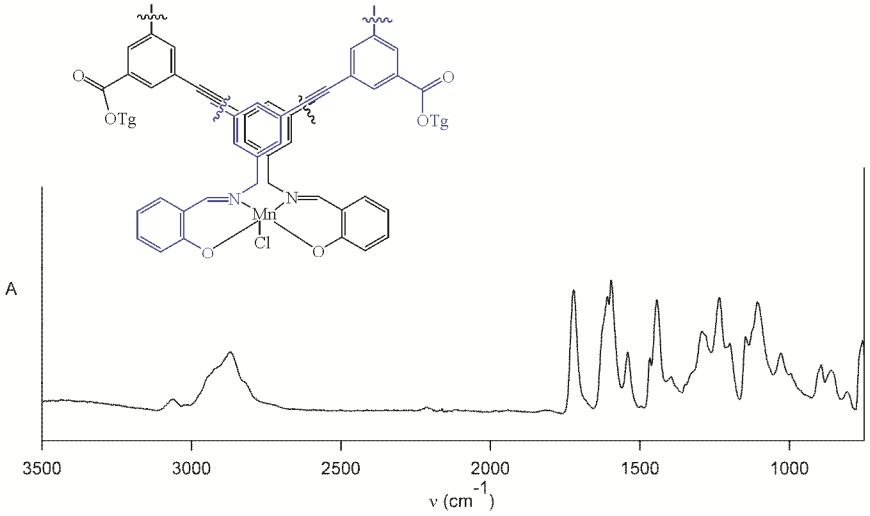


Figure G.2 FT-IR spectrum for the Mn(salen)-mPPE salen complex.